NATIONAL CANCER INSTITUTE

Center for Strategic Scientific Initiatives

THE PROVOCATIVE QUESTIONS INITIATIVE PROGRAM EVALUATION

January, 2016

NIH NATIONAL CANCER INSTITUTE Center for Strategic Scientific Initiatives

Clarivate Analytics | Formerly the IP & Science business of Thomson Reuters

Contact details

Vincent Huang (vincent.huang@thomsonreuters.com)

Scientific Analyst
Clarivate Analytics
Formerly the Intellectual Property & Science business of Thomson Reuters

TABLE OF CONTENTS

List of Tables

List of Figures

1 EXECUTIVE SUMMARY

Since 2011, the Provocative Questions (PQ) Initiative has provided support for Cancer research that addresses important questions that are broadly considered challenging or understudied. The PQ Initiative focuses on asking difficult questions, and has been implemented as an effort to solicit new approaches from diverse scientific disciplines. In 2016, an evaluation of the program was performed by Clarivate Analytics (formerly the IP & Science business of Thomson Reuters). This evaluation builds upon a previous evaluation conducted by Clarivate Analytics in 2014, and used both quantitative and qualitative methods to evaluate the PQ Initiative, with a focus on three following evaluation questions:

1. How effective are PQ program processes?
2. Did the size of PQ research areas increase following the issuance of each PQ?
3. Has the PQ initiative supported high quality and novel science in the targeted areas?

The key findings of the evaluation are summarized below.

How effective are PQ program processes?

Overall, the PQ program processes were found to be effective. Interviewees generally found the question development process to be democratic and inclusive. In addition, the members of the Executive Committee and the Workshop Participants indicated that workshops were a useful mechanism to develop and select questions, despite the high level of effort required to administer the PQ program. The initiative has successfully targeted research areas that are underrepresented in the overall Cancer[1] field, with the majority of PQs (91%) targeting research areas that were not well represented in the literature prior to the issuance of a question. The requirement to retire PQs was well received by Applicants and Awardees, and was seen as an important feature of the program that ensures that the most current PQs are the focus of the funding initiative. However, there was disagreement over which specific questions should have been retired. The Reviewers and other stakeholders interviewed recognized that the PQ initiative is a unique mechanism and has distinctive requirements. The majority of both groups recognized the lessened emphasis on preliminary data in the Request for Application (RFAs) as a strength; however, not all participants in the program were aware of that fact prior to applying or serving as a Reviewer. An evaluation of the scores received by the applications during the review process revealed that the Reviewers took this de-emphasis into account.

Did the size of PQ research areas increase following the issuance of each PQ?

One of the overarching goals of the PQ program is to highlight understudied and important Cancer research areas. There was an increase in the estimated total share of Cancer research in two thirds of the PQs after each question was incorporated into the program. Additionally, the estimated numbers of authors working in the research areas increased for 85% of the PQs after the questions were introduced. It is important to note that there may be other factors beyond the PQ program that are causative for this increase (such as other funders or organic growth of a research field), particularly as the PQ program funds a small percentage of the publications in most of PQ research areas. An analysis of all publications supported by National Cancer Institute (NCI), indicated that the majority of PQ research areas are attracting less new Principal Investigators (PIs) to NCI than the average NCI R01 and R21 grants, suggesting that these research areas may not be as attractive to researchers starting their independent careers and are not an easy entry point into the Cancer research field for investigators that have an alternative research focus.

[1] We defined the Cancer field of research as a corpus of publications selected by a combination of business rules and keywords as outlined in Appendix C.

Has the PQ initiative supported high quality and novel science in the targeted areas?

A preliminary evaluation of the impact of the PQ program was performed. On average, each PQ funded project has produced four publications. Approximately half of these publications were in the research area associated with the PQ through which the project received funding. Since the PQ portfolio intentionally targeted areas with more risk associated with them, one would not expect all publications to be in the original targeted research area. The normalized citation impact of papers funded by the PQ program is twice as high as the other papers in the PQ research areas. The majority of Awardees (85%) indicated that they had new research findings that directly resulted from PQ funding. In addition, 65% of Awardees had identified new methods or model sets. Branch Chiefs and Program Directors cited promising approaches and early successes. These included work in cachexia; social and neuroscience advances in message processing; the role of positive emotions in physical exercise; and biological aging and colon cancer.

2 PROGRAM OVERVIEW AND EVALUATION GOALS

This report provides an overview of an evaluation of the National Cancer Institute's (NCI's) Provocative Questions (PQ) Initiative, conducted in 2016 by Clarivate Analytics. In this section of the report, a brief introduction to the PQ program is provided, as well as a description of the three main evaluation questions used to guide the evaluation. Finally, an overview of the evaluation approach is provided.

2.1 PROGRAM OVERVIEW

Initiated in fiscal year (FY) 2011, the PQ Initiative provides support for Cancer research that addresses important questions that are broadly considered challenging or understudied. The PQ Initiative complements the NCI's broader funding portfolio with a more flexible Request for Application (RFA) design, a focus on asking difficult questions, and an effort to solicit new approaches from diverse scientific disciplines. The PQ Initiative has solicited applications in 2011, 2012, 2013, and 2015 (the most current issuance).[2] At the time of this evaluation, additional applications for the 2015 issuance were still being reviewed; thus, this evaluation only includes the 2011, 2012, and 2013 issuances. In addition, the 2013 issuance was only included in the qualitative component of this evaluation. This issuance was not included in the quantitative component of the evaluation because it was deemed too early to observe any sizable effect of the PQs that were newly implemented in this issuance.[3]

In its first three issuances, the PQ Initiative funded 10.3% of applications submitted (188/1822), for a total of $183.2 million in new awards (**Table 2.1**).[4] The PQ program utilized two funding mechanisms: R21 and R01. To date, the majority of grant applications (58%) and awards (62.2%) have been R01s.

Table 2.1 – Summary of PQ Initiative applications, awards and total cost

	R01			R21			ALL:R01 & R21			
Issuance	Applications	Awards	Funded Amt.	Applications	Awards	Funded Amt.	Applications	Awards	% Awarded	Funded Amt.
2011	422	38	$66.0M	332	18	$6.7M	754	56	7.43%	$72.7M
2012	460	59	$73.8M	317	35	$13.2M	777	94	12.10%	$87.0M
2013	170	20	$17.3M	121	18	$6.1M	291	38	13.06%	$23.5M
Total	1052	117	$157.1M	770	71	$26M	1822	188	10.3%	$183.2M

2.2 EVALUATION GOALS

This evaluation focuses on the following three overarching evaluation questions, and when appropriate, considers the impact in the targeted research areas before and after the establishment of the program:[5]

[2] There was no issuance in 2014.
[3] The median lag to publication for a new grant at National Institutes of Health (NIH) is 3 years.
Boyack KW, Jordan P. (2011). Metrics associated with NIH funding: a high-level view. Journal of the American Medical Informatics Association 18:423-431
[4] This amount includes all years of funding. (R01: four years, R21: two years)

1. **Are PQ program processes effective?**
2. **Did the size of PQ research areas increase following the issuance of each PQ?**
3. **Has the PQ initiative supported high quality and novel science in the targeted areas?**

Qualitative and quantitative results will be provided for each of the evaluation questions above. The report will focus on program wide trends. A productivity analysis, broken down by subtopics deemed relevant to each PQ, is provided in **Appendix G**.

2.3 EVALUATION METHOD

A mix of literature-based quantitative and feedback-based qualitative evaluation approaches were utilized throughout the evaluation. Feedback was provided throughout the evaluation process by an independent Evaluation Advisory Committee. The evaluation was conducted using methods that included topic modeling, a bibliometric analysis of the relevant literature, as well as interviews and surveys of key stakeholders.

Feedback from various stakeholder groups was solicited in the spring of 2016 using interviews and survey instruments (**Appendix B**). Interview and survey questions were focused on the PQ development process, application process, review process, management process, quality, and the scientific outcome of PQ research. In addition, the stakeholders were also asked about their perception of community enthusiasm, innovation of the projects and successes of the PQ program. Key stakeholder groups and participants interviewed included Executive Committee members, Branch Chiefs/Program Directors, Reviewers, and Workshop Participants. In addition, an online survey was conducted to solicit feedback from the PQ program Applicants (whose applications were not awarded) and Awardees. The Science and Technology Policy Institute designed and conducted focus groups with stake holder representatives organized in late 2015 to inform the design of the interview and survey guide. The interview and survey guides were designed by the Madrillon Group and vetted by the Evaluation Advisory Committee.[6] The interviews and surveys were conducted by the Madrillon Group.

A quantitative analysis based on proposal applications, funded publications, and PQ-related literature was also conducted. In this evaluation, an explicit assumption made is that the number of PQ-related publications can be used as a proxy for the size of each PQ research area. The quantitative analysis targeted 33 Provocative Questions (questions for short) from the 2011 and 2012 issuances and looked at literature published between 2008 and 2015. Since it would be impractical to review all scientific publications for inclusion in the evaluation, a mixture of machine learning and subject matter expert (SME) review techniques were employed to identify PQ-related publications. Business rules were developed in a previous evaluation to identify Cancer publications between 2008 and 2015. Topic modeling was used as a practical tool to assist NCI SMEs with the identification of PQ-related publications from the approximately 363,000 Cancer publications identified during this time period. Importantly, the NCI SMEs were relied upon to interpret the nuances of each PQ to isolate PQ-related publications from this candidate set for subsequent analyses. The resultant sizes of PQs in terms of publication count between 2008 and 2015 range from tens to thousands of papers. More detailed procedures are described in **Appendix C**. Project information and reported funded publications were identified and obtained from the NIH

[5] For each PQ, the "Before PQ" period is defined as the four years before the first appearance of the Question in a RFA and includes the year that the RFA was issued. The "After PQ" time period starts the year after the RFA is issued until the end of 2015.

[6] The surveys and interview guides were cleared for use by the NCI Office of Management and Budget in February 2016.

RePORTER system. Additional application-specific data such as criterion and overall impact scores of applications were obtained via the Query, View and Report (QVR) system of NIH.

3 KEY FINDINGS

3.1 ARE PQ PROGRAM PROCESSES EFFECTIVE?

The PQ Initiative was implemented with atypical processes to facilitate its unique programmatic goals. Here the following areas are evaluated to understand the effectiveness of the overall program: the question development process, how well the questions targeted important but understudied areas, how the question retirement process was perceived by the research community, how successfully the RFA described the unique PQ requirements to the research community, and whether these unique requirements were taken into account when applications were reviewed.

3.1.1 Question development process

Despite the high level of effort required to administer the PQ program, interview and survey responses indicate that the question development process was democratic, inclusive and of high-quality. However, there is room for improvement.

The PQ question selection process involves community outreach, solicitation of ideas from workshops, and reviews of the candidate questions by the PQ Executive Committee. Interviews of key stakeholders and surveys of Applicants and program Awardees were used to assess the strengths and weaknesses of the question selection process. This evaluation was in part motivated because of the extra time necessary to implement the PQ program over a more traditional funding approach at the NCI, a point that was noted during many of the interviews with the Executive Committee, Branch Chiefs, and Program Directors.

Interviewees from all groups (Executive Committee, Branch Chiefs, Program Directors, Workshop Participants, and Reviewers) generally found **the question development process to be democratic and inclusive.**[7] In addition, the Interviewees found that **workshops were a useful mechanism to develop and select questions**. Many of those interviewed, both within NCI and in the extramural community, applauded the inclusive process used to develop both the PQ initiative and the questions. Interviewees specifically noted the value of including new and diverse cohorts of individuals in the workshops to broaden perspectives. Many Interviewees commented that the question development process had improved over time, as the number of questions was whittled down to a more manageable number. Many Interviewees also said that the resulting questions were well-worded. However, some Workshop Participants and Reviewers interviewed noted that not everyone external to NCI understood the uniqueness of the Initiative and the need to conceptualize their work to develop questions during PQ workshops, or to review PQ grants using a different lens.

Interviewees noted several areas for potential improvement. Approximately half of the interviewed Branch Chiefs and Program Directors indicated there is a need for fewer questions that are better articulated. For example, one Interviewee recommended focusing in on the 10 to 12 questions that are the most critical. Another Interviewee noted that asking the research community to rank questions would both screen for scientific validity and provide a ranking of importance. The Interviewees also encouraged the program to continue to include a diverse group of individuals in future workshops and to provide Workshop Participants with feedback after the process is complete. Beyond the workshop process, the Interviewees identified other ways that NCI could solicit input such as reaching out through existing means (national organizations, annual meetings, existing journals, and/or U mechanisms). In terms of alternatives to the online question submission that NCI had attempted, a few Workshop Participants mentioned the prospect of using web-based forums or alternatively, convening more

[7] It should be noted that since we did not include those who did not participate in the question development process, that this assessment could be biased.

regional workshops. Additionally, a Workshop Participant suggested that the PQ program staff review the burden required to participate in the existing online question submission process.

Survey participants were also asked about the question development process and responses differed between Applicants and Awardees. While unfunded Applicants found the development process to be less inclusive, the Awardees found the development process to be inclusive because it provided the Cancer research community with the opportunity to voice their opinions during the development process. Only 40% of Applicants, but 88% of Awardees agreed or strongly agreed that "*the PQ process gave members of the Cancer research community an opportunity to voice their opinions.*" In addition, 42% of Applicants and 60% of Awardees who also participated in the PQ development process felt that their input was reflected in the final selection of questions. There was closer alignment between the Applicants and Awardees regarding the scope of the questions; 75% of Applicants and 97% of Awardees felt that the scope of the questions were appropriate.

3.1.2 PQ targeted research areas

Important and under-represented Cancer research areas were targeted.

A key goal of the PQ program is to target areas of cancer that are challenging and understudied. Therefore, an appropriate measure of the success of the PQ selection process is the degree to which the questions collectively target under-represented areas in Cancer research. In order to determine if the selected PQs' collectively targeted areas that were understudied, the fraction of Cancer publications in each PQ topic area was estimated prior to the first issuance of the relevant PQ.[8] Thirty out of 33 PQs (91%) and their targeted areas represented less than 1% of the Cancer research field before the start of the PQ Initiative. Two thirds (67%) of the PQs targeted research areas represented by less than 0.2% of contemporaneous Cancer publications (**Figure 3.1** and **Appendix D**). Clearly, **the majority of PQs targeted research areas were not well represented in the Cancer literature** prior to the issuance of the question.

This conclusion is supported by the results of a survey of both PQ Applicants and Awardees. Ninety seven percent of Applicants and 90% of Awardees agreed with the statement "*the Provocative Questions represent perplexing, difficult to address, paradoxical, or understudied areas that should be investigated.*" The participants in the survey represent researchers with close ties to the PQ research areas. Therefore, these researchers can be assumed to have close knowledge of the PQ topic areas and how well represented the questions are in the overall field of Cancer research. In addition, 97% of Applicants and 76% of Awardees agreed that the developed PQs "*truly address gaps in Cancer research.*" The Interviewees agreed with this assessment, noting that the selected **PQ questions represent perplexing, understudied areas.**

[8] For method details, see Appendix C

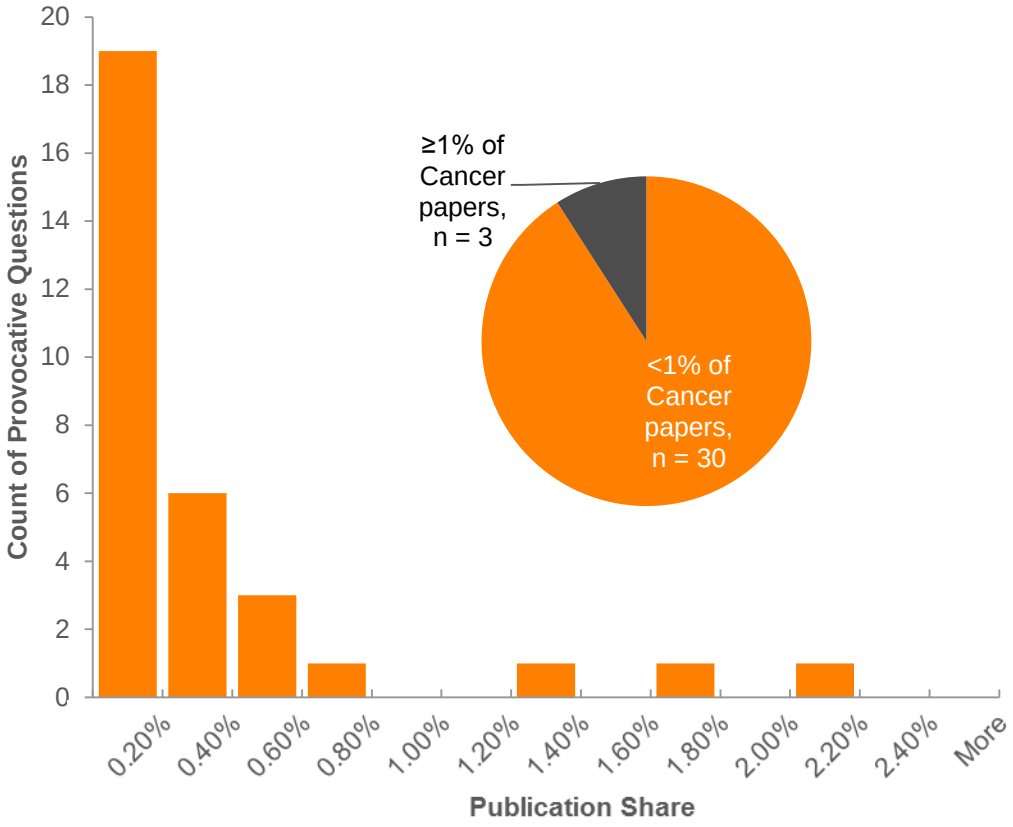


Figure 3.1 – PQs categorized by the share of Cancer publications related to the PQ research area prior to question issuance.

The bar chart shows the distribution of PQs in terms of the overall publication share of Cancer publications associated with each PQ research area from 2009 to question issuance, prior to the first issuance of each question. The pie chart shows that 30 out of 33 (91%) of all PQs represented less than 1% of Cancer literature where n is the count of each individual PQ.

3.1.3 Question retirement process

The process of updating, renewing, and retiring PQs was found to be an important feature of the program and ensures that the most current PQs are the focus of the funding initiative.

A number of PQs have been retired during the history of the program. When a question is retired, it is not rewritten or included in the next issuance of the RFA. The majority of Applicants and Awardees agreed that **the process of updating, renewing, and retiring PQs is an important feature of the program and ensures that the most current PQs are the focus of the funding initiative.** Almost all (99%) of Awardees and 85.6% of Applicants agreed with the following statement: "*The process of updating, renewing and retiring provocative questions is an important feature of the Provocative Questions Initiative.*" In addition, 96% of the Awardees and 74.2% of Applicants agreed with the following statement: "*The process of updating, renewing and retiring provocative questions ensures that the most currently provocative questions are the focus of the PQ funding initiative.*"

Applicants and Awardees had different opinions concerning if questions were retired prematurely. Thirty-one percent of Applicants believed that questions had been prematurely retired. However, 61% of Awardees believed that a question had been prematurely retired. There are several factors to consider when interpreting this difference. For example, Awardees working on a question potentially benefit if that particular question continues

to be part of the program and may prioritize areas for funding that are close to their own research area. The two groups did agree on the top three questions that were retired too soon: 6, 8 and 1 (**Table 3.1**). The research area associated with PQ 1 ("obesity and cancer") did see a significant increase (33.7%) in the fraction of Cancer publications associated with the topic after the issuance of the question (**Appendix D**)[9]. The research areas associated with PQ 6 and 8 did not see a significant change in the fraction of publications associated with the topic pre and post issuance of the question.

Table 3.1 – Top three questions from each survey group considered to be prematurely retired.

PQ	PRECIS	QUESTION	RANK BY AWARDEES	RANK BY APPLICANTS
6	disease correlation	What are the molecular and cellular mechanisms by which patients with certain chronic diseases have increased or decreased risks for developing cancer, and can these connections be exploited to develop novel preventive or therapeutic strategies?	2	1
8	tissue-dependent phenotypes	Why do certain mutational events promote cancer phenotypes in some tissues and not in others?	1	2
1	obesity and cancer	How does obesity contribute to cancer risk?	3	3

3.1.4 Clarity and design of RFA

The RFA guidelines and lessened emphasis on preliminary data encouraged Applicants to propose novel and innovative approaches.

The majority of Awardees (89%) and Applicants (66%) agreed that the RFA guidelines led them to propose novel or innovate approaches that they might not otherwise have proposed to NCI. In addition, **the majority of Awardees (67%) believed that the lessened emphasis on preliminary data allowed them to propose novel or innovative concepts and approaches.** The survey results indicated that 25% of the Applicants were not aware of the lessened emphasis on preliminary data. This agrees with interview results where two-thirds of the interviewed Reviewers said that some Applicants did not take advantage of the lessened preliminary data requirements in their applications.

The majority of both Applicants and Awardees found the application materials clear and that the amount of effort required to apply to the program was similar to other funding programs. Two thirds of the Awardees believed that the "*funding period provides sufficient time to achieve the research goals.*"

3.1.5 Review process

Reviewers recognized that the PQ Initiative was a unique funding approach with specialized expectations; they placed a lessened emphasis on approach and increased the emphasis on innovation in overall scoring of grant applications.

[9] This was determined by a log likelihood ratio test. Please see **Appendix D** for details.

The relationship between the application criterion scores and the overall impact score was evaluated to gain insight into the outcome off the review process.[10] While the Approach review criterion[11] was still the most important predictor of the overall impact score in the review process[12], its influence in the scoring of PQ applications is lessened while the influence of other criteria such as Innovation and Significance increased. This was revealed by the significant correlation of the impact score and the individual review criterion of the PQ program when compared to the typical R01 and R21 applications submitted to NCI (**Figure 3.2**).[13] **There is also a statistically significant reduction of the predictive power of the Approach score to the priority score suggesting that a de-emphasis on preliminary data was taken into account by the Reviewers (Appendix E).**[14]

While most Reviewers recognize the distinctive requirements of the PQ Initiative, those interviewed said that it was still hard for some Reviewers to make the "paradigm shift" required to evaluate the applications differently than those of traditional grant mechanisms. Across interview groups, some Interviewees said that they believe there is a bias at NIH to fund and encourage safe, incremental, basic science, further demonstrating the importance of an alternative funding path for important but potentially risky projects such as the PQs Initiative.

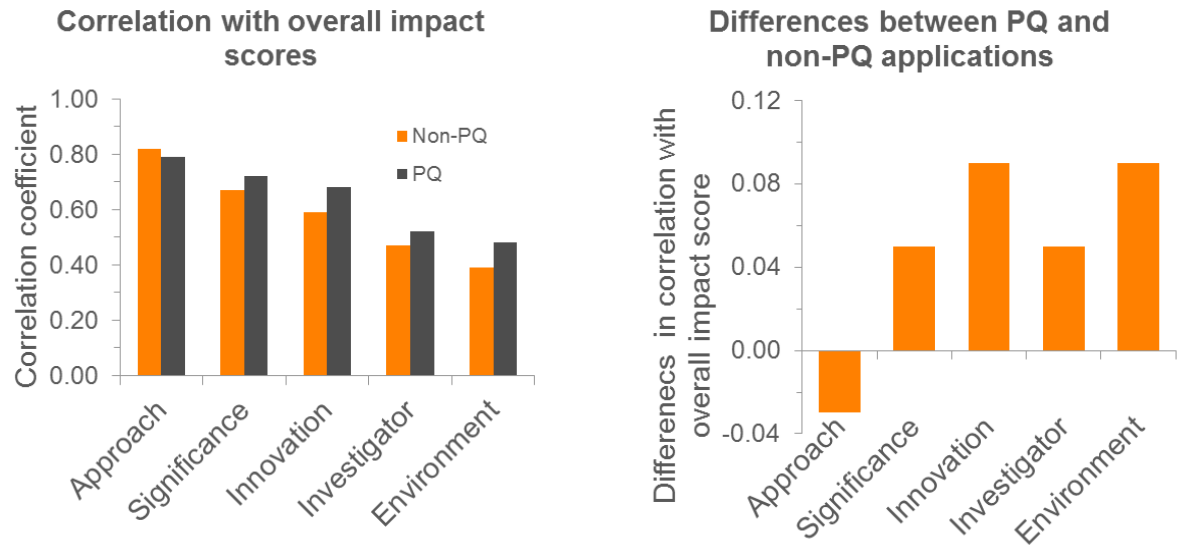


Figure 3.2 –Criterion scores and overall impact scores correlate differently for PQ and non-PQ applications

LEFT: The correlation coefficient between criterion scores and overall impact scores for PQ and non-PQ applications received between 2011 and 2014 were computed. RIGHT: The differences in the correlation between PQ and non-PQ applications.

More Awardees than Applicants agreed that the RFA led them to propose novel or innovative approaches. However, **Reviewers were split in their assessments on whether the proposals received were truly**

[10] Applications to PQ were reviewed by dedicated Special Emphasis Panels.
[11] NIH grant applications are considered in five standard criteria: Approach, Significance, Investigator(s), Innovation, and Environment. The criterion scores are given independently of the overall impact score, which is used for funding decisions.
[12] Eblen et al. (2016) How Criterion Scores Predict the Overall Impact Score and Funding Outcomes for National Institutes of Health Peer-Reviewed Applications. PLoS ONE 11(6): e0155060. doi:10.1371/journal.pone.0155060
[13] The correlation coefficient, r, ranges from 0.40 (Environment - Overall Impact Score) to 0.82 (Approach – Overall Impact Score). DOF=17,374, $p<0.0001$ for all individual criteria.
[14] Linear regression model, $R^2=0.72$, $F(20,17355)=2279$, $p<0.0001$. Please see Appendix for details.

innovative, high risk, or novel. When nine Reviewers involved in the PQ review process were individually asked "*in your opinion, were the Provocative Questions Applicants proposing innovative, high risk, or novel research*?" four respondents indicated that many Applicants were not. This sentiment appeared to be not limited to those who were interviewed and was reflected by the significantly higher (worse) criterion scores in Significance and Innovation given to the PQ applications when compared to non-PQ NCI applications (**Figure 3.3**, left panel).[15] When the same nine Reviewers were asked "*did you notice a difference in the degree of innovation, risk or novelty of the research proposed in Provocative Questions R21s compared with Provocative Questions R01s?*" again five responded no. Again, this sentiment appeared to be reflected by the significantly higher (worse) criterion scores in Significance and Innovation given to the PQ R21 applications when compared to PQ R01 applications (**Figure 3.3**, right panel).[16]

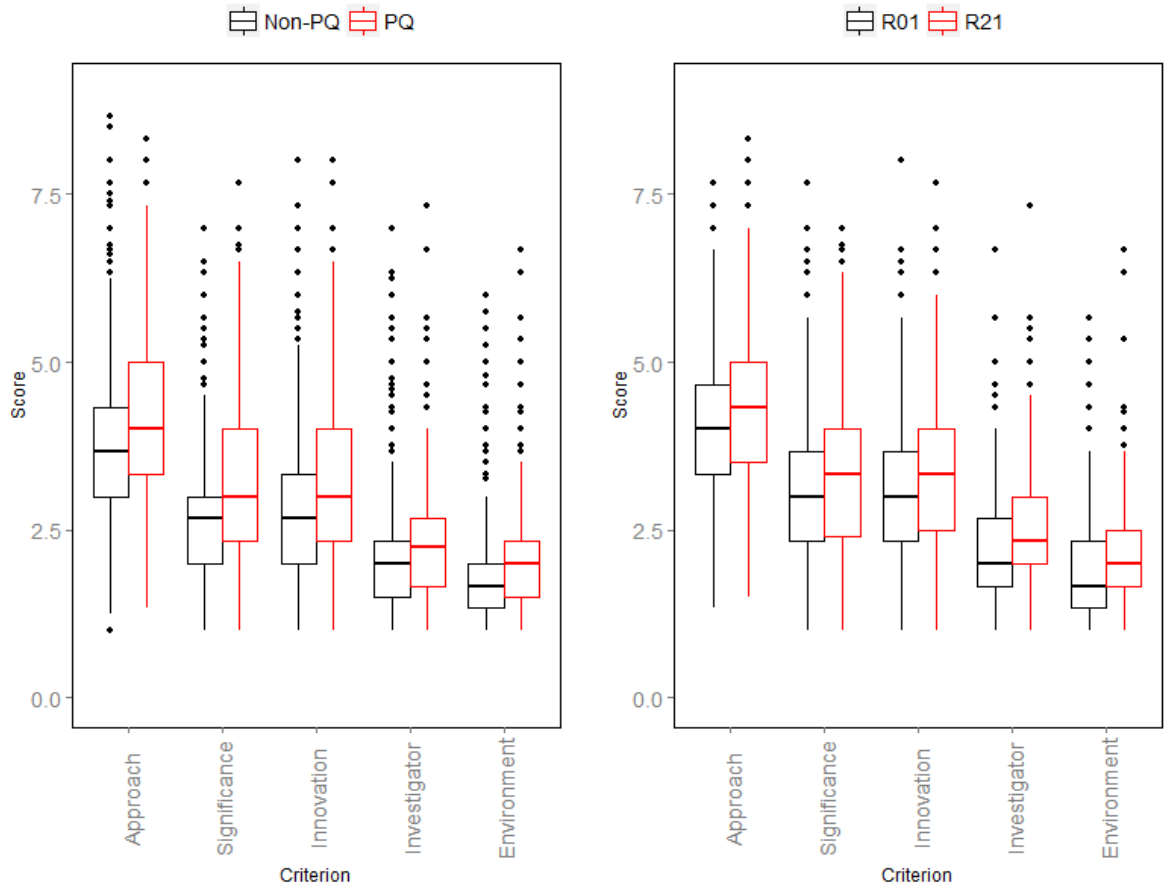


Figure 3.3 – Criterion scores received by non-PQ NCI and PQ applications between 2011 and 2014.

LEFT: Criterion scores for PQ and non-PQ applications. RIGHT: R01 and R21 applications submitted to PQ.

[15] MANOVA test, $F(5, 33028)=15.663$, $p<0.0001$
[16] MANOVA test, $F(5, 1437)=10.559$, $p<0.0001$

3.2 DID THE SIZE OF THE PQ RESEARCH AREAS INCREASE FOLLOWING THE ISSUANCE OF EACH PQ?

A key motivation of the PQ Initiative is to draw research attention to and increase research activities in areas targeted by the individual questions. In this section of the report, the level of research activity is approximated by the amount of publications in these areas and their share in the overall Cancer research landscape. This section of the report also looks at the number of new PIs in the question areas at NCI during the evaluation timeframe to understand if these research areas are drawing researches from other fields into the targeted PQ research areas. Acknowledgement of funding from other NIH institutes in the PQ research areas was also evaluated to understand what other NIH institutes and centers may have overlapping interests in funding these targeted areas.

3.2.1 Volume of publications

The majority of research areas targeted by a PQ saw an increase in that area's overall representation in the Cancer literature.

The overall volume of Cancer literature has grown linearly over the years. We computed publication share to attempt to account for the baseline increase of publications. Approximately two thirds of the research areas targeted by a PQ (21 out of 33 in total) saw an increase in the percentage share of Cancer papers published in their respective reseach area after the first issuance of that specific PQ (Appendix D). Out of the 33 PQs studied, the share size changes in 22 Questions were statistically significant.[9] Of those 22, 15 PQs (65%) had a significant increase and eight (35%) had a significant decrease (Appendix D). Two questions saw more than a 100% increase in research volume after the PQ funding question was released. The 258% increase shown in **Figure 3.4** corresponds to PQ 21 ("therapy resistance"), and the 118% increase corresponds to PQ 33 ("cachexia"). Both of these PQs targeted research areas representing a small fraction of the overall corpus of Cancer publications when the publications were first issued. The research area associated with PQ 21 (therapy resistance) was represented by 0.03% of Cancer publications prior to the PQs issuance, and the research area associated with PQ 33 was associated with 0.02% of Cancer publications prior to the PQs issuance.

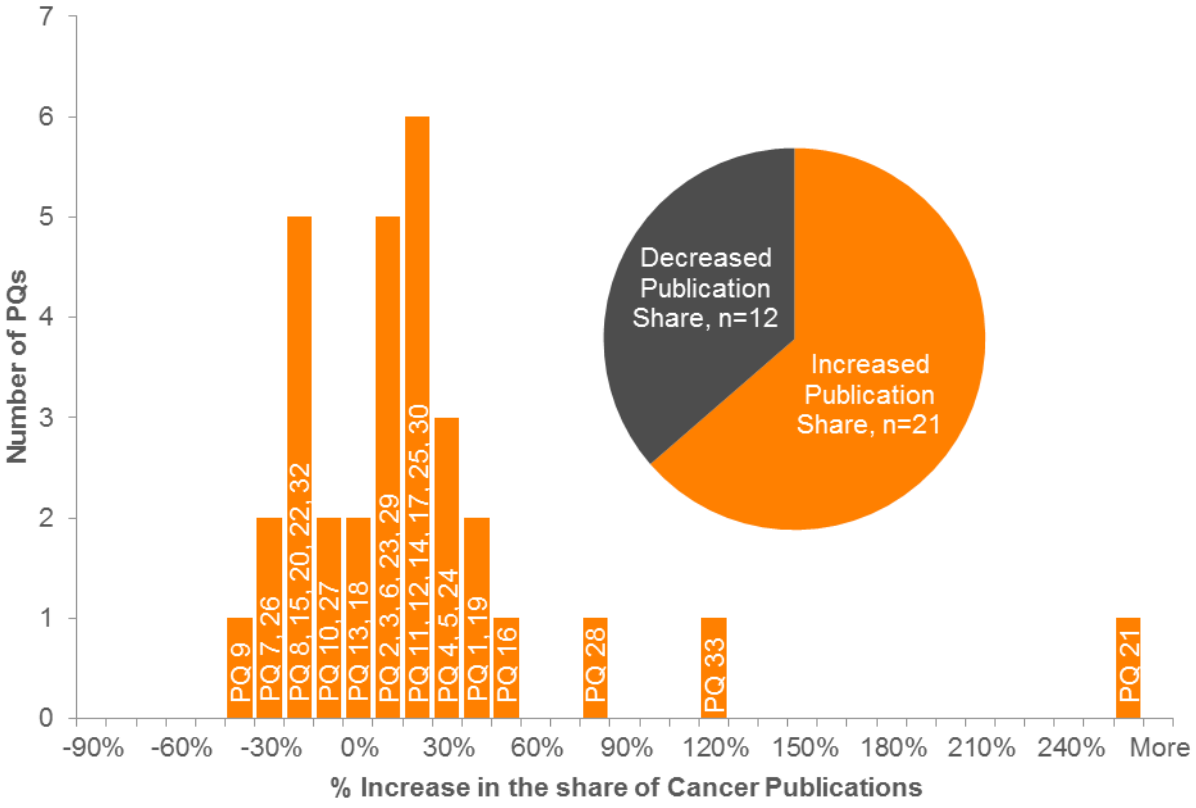


Figure 3.4 – The distribution of the percent increase in publication volume for the 33 PQs.

The bar chart shows the distribution of changes in the share of Cancer publications and the PQ numbers are listed within the bars. The pie chart insert shows that two-thirds (n=21) of PQs saw an increase in their relative share of Cancer publications.

The majority of research areas targeted by a PQ saw an increase in the number of active authors in the PQ area.[17] Eighty-five percent of PQs (and in turn their targeted research areas) saw an increase in the number of authors publishing in that research area after the PQ question was identified. For the majority of questions (87%), increases and decreases in the number of authors corresponded to increases and decreases in the share of publications in the research areas associated with the PQ program. The largest increases in the number of authors occurred in the research area associated with PQ 21 (395%) and PQ 33 (150%); these two areas also saw the largest increase in the number of publications after the issuance of the PQ program (**Appendix D**). In general, there were more authors per paper (approximately one more author) after PQ issuances, suggesting a potential increase in collaboration, and/or manpower, was needed to conduct the research. This is in alignment with recent trends in the research community, where collaboration and team science have increased during the same time period.[18]

[17] Distinct authors' names were compiled across all PQ-related publications and counted for each PQ to estimate the number of active authors in each PQ targeted research area. It should be noted that a full disambiguation of the author names was not performed. For example, "Robert Johnson" and "Rob Johnson" may count as two authors while in actuality they are the same person. Thus the absolute number of authors identified here is likely to be an over-estimate. The direction of change, however, should remain the same.

[18] https://www.nlm.nih.gov/bsd/authors1.html

3.2.2 NIH funding context

Most of the PIs in PQ targeted research areas are not new to the National Cancer Institute.

Overall NIH funding in the PQ targeted areas appears to have increased based on the acknowledgements of publications in the research areas.

The number of new NCI PIs in each PQ research area was estimated and compared to the overall number of new PIs at NCI to understand how these research areas are attracting investigators to NCI. This analysis includes all investigators publishing in the research areas who acknowledge NCI funding, not just those directly funded by the PQ program. New NCI PIs for the purpose of the evaluation are defined as investigators who had no other NCI funding between 2001 and the first issuance of the PQs RFA (2011).[19] **The average percent of new PIs across all research areas is 11.55%, lower than the percent of new PIs associated with NCI as a whole (14.39%) (Appendix F).** However, the topic areas of five PQs were particularly effective at attracting new PIs to NCI. In these five questions, 20% or more of the investigators were new to NCI, well above the NCI average. One particular standout is PQ 8, tissue-dependent phenotypes; 40% of the PIs funded in the research area associated with PQ 8 were new to the NCI.

[19] Names of the same PIs may have variants and were not disambiguated. Thus the numbers of PIs should be considered estimates.

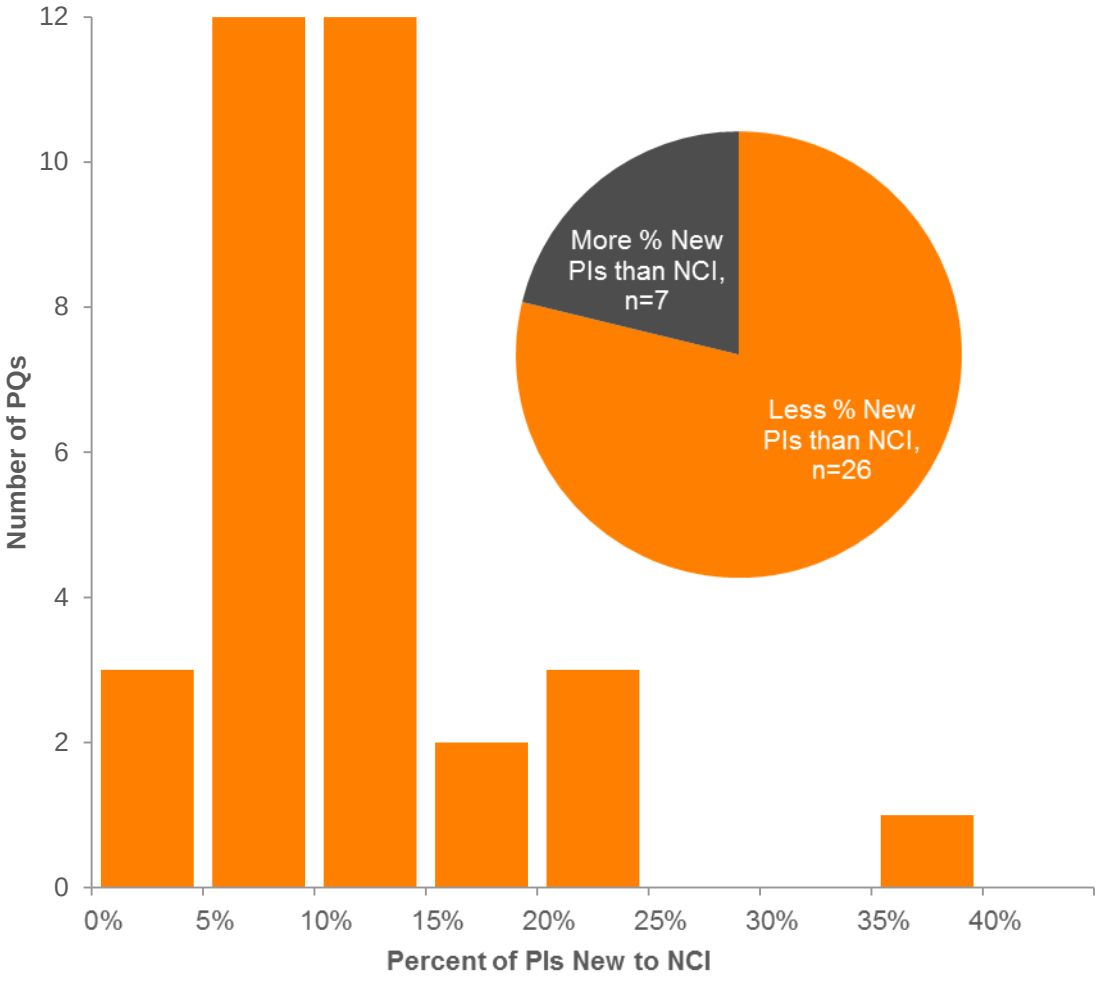


Figure 3.5 – Number of PIs new to NCI

The bar chart shows the distribution of PQs in terms of the percent of PIs that were new to NCI. The NCI average was 14.4%. The majority of PQs did not attract more investigators to the NCI than the NCI average (14.4%).

NIH funding of the research areas was evaluated in order to assess overall changes in various Institutes' funding before and after the issuance of the PQ program. In general, NIH funding in the PQ areas increased after the first issuance of the PQ program. The number of publications acknowledging NIH grants in the PQ areas, including those funded by the PQ Initiative, increased overall when comparing the four years immediately prior to the issuance of each PQ (7,540) and the four years immediately after (12,799). The top ten Institute and Centers (ICs) funding in the PQ areas (based on the number of post PQ papers) are shown in **Table 3.2**.

Table 3.2 – Top 10 NIH institutes funding research in the PQ areas ranked by the number of post PQ acknowledgements of NIH funding by publications in the research topics.

INSTITUTE & CENTERS	BEFORE PQ	AFTER PQ	% CHANGE
NIH National Cancer Institute (NCI)	3095	3779	22.1%
NIH National Institute of General Medical Sciences (NIGMS)	283	544	92.2%
NIH National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)	252	323	28.2%
National Center for Research Resources (NCRR) (dissolved 12/2011)	280	295	5.4%
NIH National Institute of Environmental Health Sciences (NIEHS)	290	268	-7.6%
NIH National Heart, Lung and Blood Institute (NHLBI)	157	218	38.9%
NIH National Institute of Allergy and Infectious Diseases (NIAID)	136	213	56.6%
NIH National Institute of Neurological Disorders and Stroke (NINDS)	113	135	19.5%
NIH National Institute of Biomedical Imaging and Bioengineering (NIBIB)	80	134	67.5%
NIH National Institute on Aging (NIA)	110	115	4.5%

3.3 HAS THE PQ INITIATIVE SUPPORTED HIGH QUALITY AND NOVEL SCIENCE IN THE TARGETED AREAS?

This evaluation makes a preliminary assessment of the quality of the research output; it analyzes how well the funded research is being cited by the field, and identifies successes as reported by the program officers and Awardees. This evaluation was conducted acknowledging that significant impact may not be observable within the short time frame since the PQ Initiative was established.

3.3.1 Percent of PQ funded publications that are related to the question under which they were funded

PQ-funded projects have been productive; however, approximately half of their publications were in the research area associated with the PQ through which the project received funding.

Using NIH RePORTER, 614 papers were associated with the PQs included in this evaluation. **On average, each funded project produced approximately 4 papers,** which is in line with NCI's median of four papers during the same time period (R01 and R21 only). These 614 papers were reviewed to determine if the subject area of the paper is aligned with the research area(s) of the PQ that was the target of the original grant (see **Appendix C**). A

total of 48% of these papers (312) were deemed to be relevant to the target PQ publication area.[20] The individual questions varied in terms of the total number of publications, and the percent of papers related to the PQ topic area (**Figure 3.6**). Questions located in the top right quadrant of this figure fund projects that are highly productive (in terms of the number of publications) relative to the other PQs and produce a large percent of publications in the original target research area. Projects funded by PQs that publish less than 50% of publications in the topic area of the PQ are not necessarily indicative of low program performance because high risk projects often do not produce the results that are originally expected. This result is an indication that the overall portfolio of questions included those in higher risk/newer areas of research.

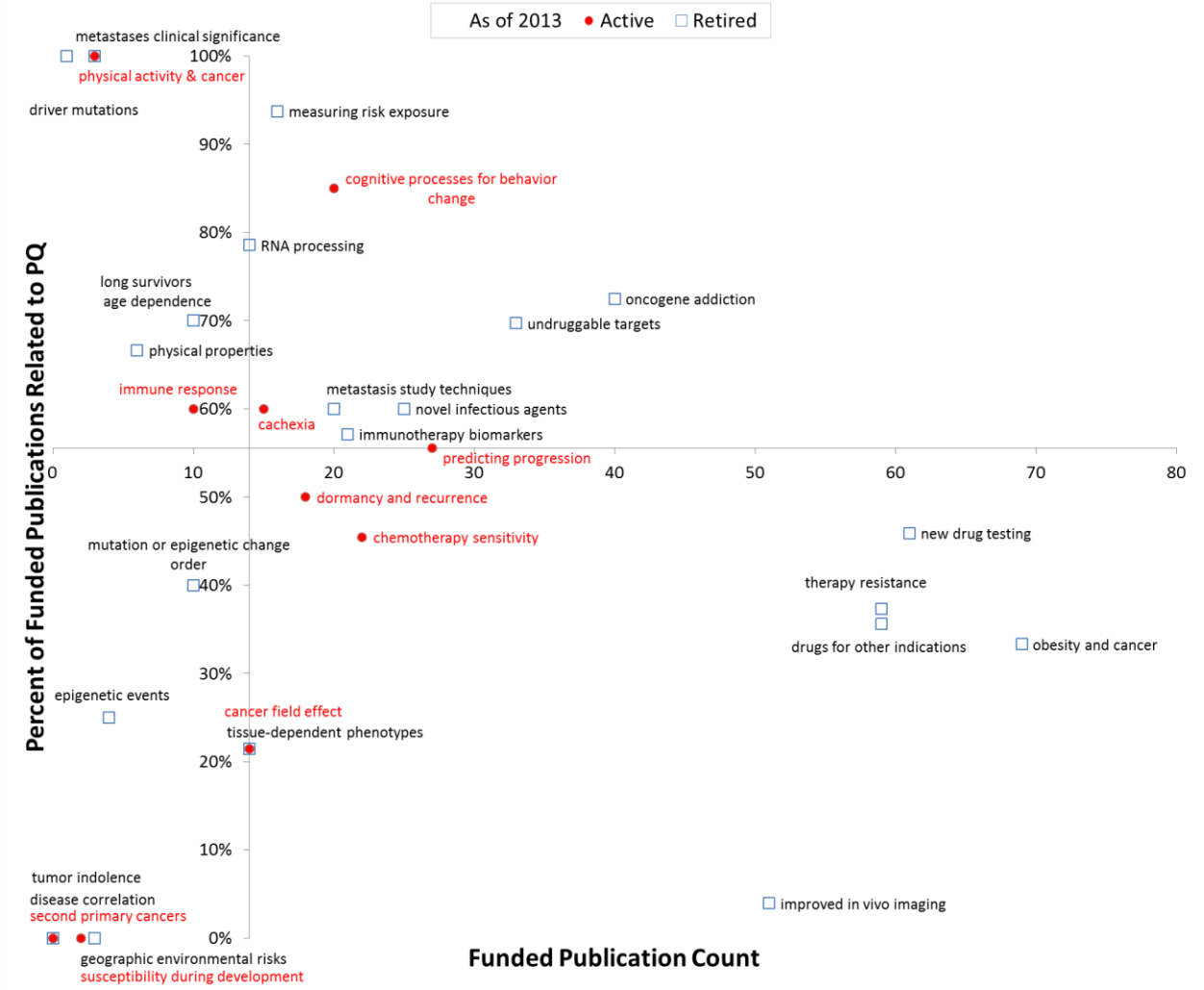


Figure 3.6 – A map of PQs based on the overall funded publications and the percent of funded publications relevant to PQ

More productive PQ's are plotted toward to the right. PQs with projects that produced more relevant publications are plotted toward the top. The axes cross at the median point of the two dimensions in the graph. The blue squares and red dots indicate retired PQs and active PQs as of the 2013 issuance. By 2015, all questions but "cancer field effect" were retired.

[20] This statistic is descriptive for the PQ Initiative and not to be compared with other programs. The same process to establish a comparative baseline for other NIH programs may be prohibitively costly in terms of manpower.

Novel work has a higher degree of risk associated with it. If the grant selection process was successful at identifying novel and innovative approaches, one would expect a proportion of the researches to obtain results that contradict their initial research hypothesis. Thirteen percent of Awardees indicated in the survey that they obtained research results that definitively contradicted the initial research hypothesis. Eighty-five percent of individuals who indicated that they encountered contradictory results believed that this was the result of the high-risk nature of their initial work.

3.3.2 Citation impact of publications supported by PQ and published within the PQ targeted research areas

PQ-funded publications are cited twice as much as other publications in the same field.

In order to determine the impact of the papers associated with each research area, the field normalized citation impact of papers was determined. The normalized citation impact is used instead of comparing directly the number of citations received by publications because publication rates (and therefore citation rates) vary between research field and with time. The normalization factor is the world average citations per paper for the year and the subject[21] in which the paper was published. This allows papers to be compared to the world average (a normalized citation impact of 1.00), and allows for comparisons between different groups of publications.

The average field-normalized citation impact is almost twice as high for PQ funded papers than for non-PQ funded papers (Figure 3.7). This difference is statistically significant.[22] Additionally, there was a small but significant increase of field normalized citation impact for non-PQ funded publications in the research areas after the initiation of the project.[23] This increase could potentially be associated with an increase in the number of individuals in a particular research area (which would increase the potential intellectual reach of a scientific work). Alternatively, this increase could be associated with the maturation of the scientific areas associated with each PQ research area. Notably, the average normalized citation impact for papers in the PQ research areas was greater than the world average (1.00). This is further evidence that the selected PQs target important research areas.

[21] In this evaluation, the Web of Science Journal Subject Categories were used as a proxy for research field.
[22] Significance of difference was tested using the two-tailed Student's t-test, t(211)=3.003, p=0.003
[23] Significance of difference was tested using the two-tailed Student's t-test, t(23493)=3.698, p=0.0002

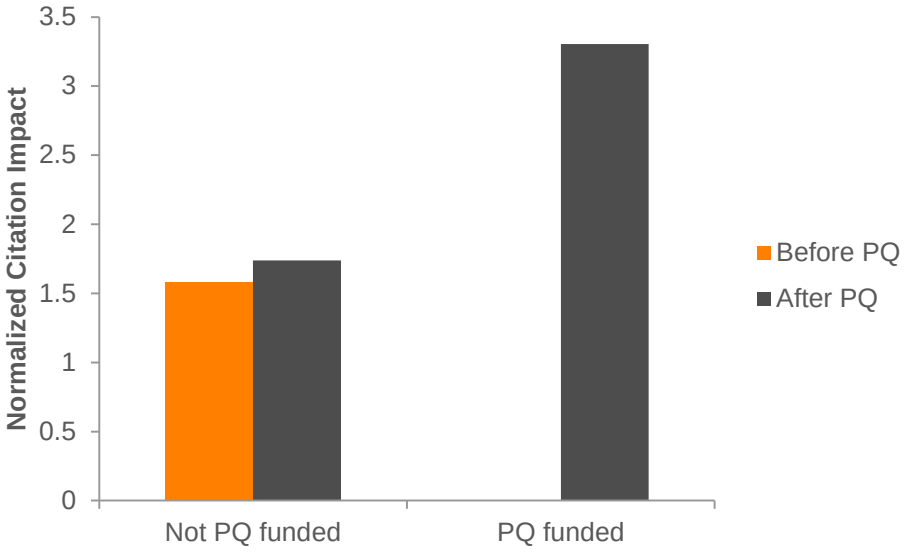


Figure 3.7 – Comparison the field-normalized citation index of the PQ and non-PQ funded publications in PQ targeted areas.

3.3.3 Survey and interview results regarding the value and impact of PQ funded research

While it is too early to understand the full impact of PQ funding, Awardees identified both short and potential long term results from their research.

The majority of Awardees (85%) indicated that they had new research findings that directly resulted from the PQ funding (**Figure 3.8**). In addition, 65% of Awardees had identified new methods or model sets. Regarding the longer term potential for research results, 72% of Awardees expected new lines of research to result from the PQ funding and that their research would lead to a changed understanding of a new disease. **Overall, the majority of Awardees expected to see both short-term and long-term results from their research.**

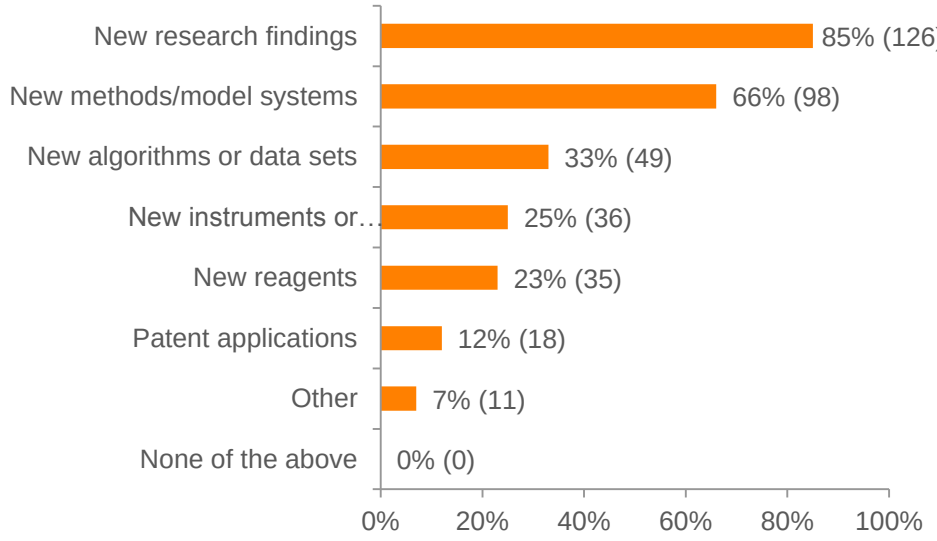


Figure 3.8 – Short-term scientific results of Awardees' research

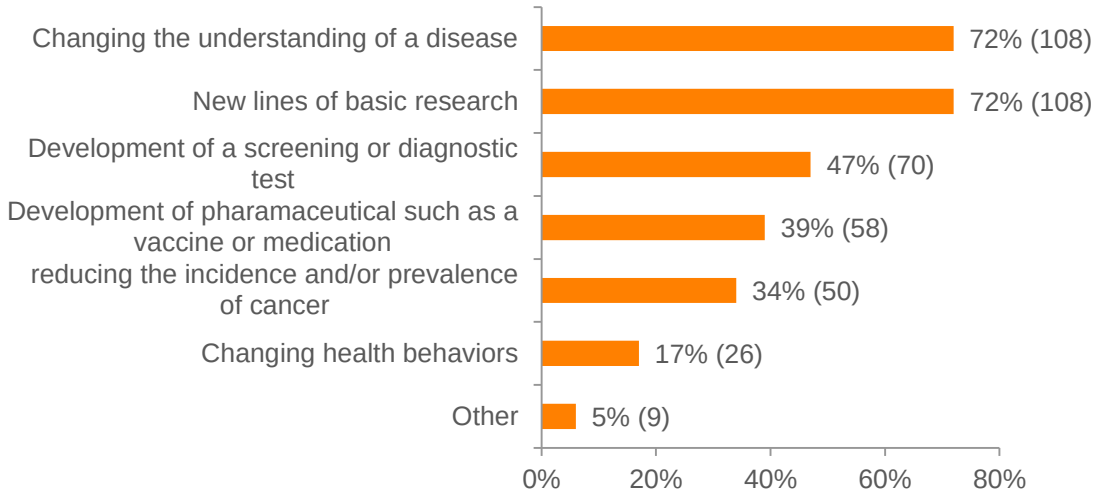


Figure 3.9 – Potential long-term results of PQ research

In addition, the Awardees were asked for specific scientific successes from their funded research. A total of 133 Awardees provided 159 specific examples of scientific successes, some of which are shown in below in no specific order.

- *"We have discovered a new molecular pathway that produces dormant, epigenetically plastic, treatment-resistant cancer cells"*
- *"We have developed a mouse model for isolating intact nuclei from tumor derived cells"*
- *"The findings that we obtained show for the first time that chemotherapeutics, such as Folfiri, contribute to develop muscle mass loss and muscle weakness by activating MAPKs (ERK1/2 and p38) activation and by promoting mitochondrial toxicity."*
- *"Through PQ funding, we have established novel imaging and sequencing modality through multidisciplinary collaborations, which allows us to study the earliest stage of metastasis progression."*
- *"We have developed a potent and widely applicable neuroblastoma model that can be used to genetically probe any aspect of the immune-tumor interface."*
- *"Development of computational methods to reconstruct the evolution of cancer even if longitudinal samples are not available."*
- *"...better methods for assessing individuals' exposure to environmental carcinogens"*
- *"The identification of communication strategies that persuade recipients to engage in cancer prevention behaviors."*

Nine percent of respondents noted that it was too early to point to specific examples of scientific successes from the grant funding.

During the interviews, the Branch Chiefs and Program Directors were also asked to articulate the value of the science funded as part of this initiative. The majority of those interviewed said that it was too soon to judge the success of the science conducted under the Initiative, with one noting that due to the high-risk, high-reward nature of the work "we should be able to tolerate failure". Still, some Branch Chiefs and Program Directors **cited promising approaches and early successes**. These included work in **cachexia, social and neuroscience advances in message processing, the role of positive emotions in physical exercise, and biological aging and colon cancer**. Interestingly many of these research areas are targeted by PQs that have a higher percentage of funded publications related to PQs, and not necessarily in areas where there have been the most PQ funded projects or most PQ funded publications (**Figure 3.6**).

4 RECOMMENDATIONS

In this evaluation, the following three evaluation questions were considered:

1. Are PQ program processes effective?
2. Did the size of PQ research areas increase following the issuance of each PQ?
3. Has the PQ initiative supported high quality and novel science in the targeted areas?

Regarding the **PQ program processes**, it was found that the overall PQ program processes were effective in targeting important research areas that are underrepresented in the overall Cancer field. Interviewees generally found the question development process to be democratic and inclusive and the workshops were a useful mechanism to develop and select questions, despite a significant amount of effort is required for the current process. While the lessened emphasis on preliminary data was recognized as a positive feature of the program and was taken into account when applications were reviewed not all Applicants and even Awardees were aware of this aspect of the Initiative. Lastly, the Applicants and Awardees were not in agreement over which specific questions should have been retired.

Given these results, it is recommended that the PQ program management:

- Continue to improve and optimize program management processes, including question development and application review to lessen burdens on participating program officers.
- Provide summary of retirement decisions to improve transparency.
- Clarify the de-emphasis on preliminary data more in future RFAs.
- Continue to track the chronological trend of the publication related to PQ in the next several years to determine whether there is an optimal number of issuances to catalyze risky or understudied research in fields, and whether there is a difference between those PQs viewed as retired too soon or too late.

Regarding the support PQ Initiative provides to **increase the size of PQ targeted research areas**, it was found that all PQs successfully targeted understudied areas and in two thirds of the PQs, there was an increase in the estimated total share of Cancer research since the establishment of the PQ Initiative, after correcting for a baseline increase in Cancer research. The NCI remained the principal force among NIH ICs in funding research in these areas. While the estimated numbers of authors have also increased, most of the PIs in PQ targeted research areas are not new to the NCI.

Given these results, it is recommended that the PQ program management:

- Review the role, experience, and type of researchers best suited to pursue such risky projects and consider strategies to target these candidates.
- If there is a desire to attract new or early career researchers to NCI in the PQ research areas, consider what might be potential barriers to new researchers. For example, an early-career researcher may look to apply for traditional R01s where there may be options to renew or re-compete at the end of the maximum funding period. As another example, R21 may not be the best option for early-career researchers as it limits the project period to 2 years.
- Investigate methods to identify and target candidate researchers who may be interested in future PQs, especially those not in NCI who are already working in related areas.

Regarding the support the PQ Initiative provides to **support high quality and novel science in the targeted areas**, it was found that PQ-funded projects have been productive and their publications have been cited almost twice as much as other publications in the same field. However, approximately half of their publications were in the research area associated with the PQ through which the project received funding. While we recognize that it may still be early for these project's impact to fully develop and for the community to fully appreciate the results, there are promising successes identified both by the funded investigators and the Program Directors.

Given these results, it is recommended that the PQ program management:

- Consider realistic expectations of research outcome for each PQ for better measurement of impact.
- Revisit the evaluation in five years after there has been sufficient time for the impact of the program to be fully developed and observable.
- Characterize research areas that have high percentage of PQ-related publications to inform PQ development and subsequent program management.
- Evaluate the type of grants and support that are needed by investigators to initiate research in PQ areas after questions are retired.

APPENDIX A LIST OF PQS

PQ	PRECIS	RFA YEAR	PQ ID IN RFA	QUESTION TEXT
1	obesity and cancer	2011	PQ1	How does obesity contribute to cancer risk?
		2012	PQA2	
2	geographic environmental risks	2011	PQ2	What environmental factors change the risk of various cancers when people move from one geographic region to another?
3	measuring risk exposure	2011	PQ3	Are there ways to objectively ascertain exposure to cancer risk using modern measurement technologies?
		2012	PQA4	As modern measurement technologies improve, are there better ways to objectively ascertain exposure to cancer risk?
4	cognitive processes for behavior change	2011	PQ4	Why don't more people alter behaviors known to increase the risk of cancers?
		2012	PQA3	How do cognitive processes such as memory and executive function interact with emotional or habitual processes to influence lifestyle behaviors and decisions, and can we use this knowledge to design strategies to change behaviors that increase cancer?
		2013	PQA1	How do decision-making processes influence habitual behaviors, and how can that knowledge be used to design strategies that lead to adoption and maintenance of behaviors that reduce cancer risk?
5	drugs for other indications	2011	PQ5	Given the evidence that some drugs commonly and chronically used for other indications, such as an anti-inflammatory drug, can protect against cancer incidence and mortality, can we determine the mechanism by which any of these drugs work?
		2012	PQA1	What is the molecular mechanism by which a drug (such as aspirin or metformin) that is chronically used for other indications protects against cancer incidence and mortality?
6	disease correlation	2011	PQ6	What are the molecular and cellular mechanisms by which patients with certain chronic diseases have increased or decreased risks for developing cancer, and can these connections be exploited to develop novel preventive or therapeutic strategies?
7	age dependence	2011	PQ7	How does the life span of an organism affect the molecular mechanisms of cancer development and can we use our deepening knowledge of aging to enhance prevention or treatment of cancer?

PQ	PRECIS	RFA YEAR	PQ ID IN RFA	QUESTION TEXT
		2012	PQB4	What mechanisms of aging, beyond the accumulation of mutations, promote or protect against cancer development?
8	tissue-dependent phenotypes	2011	PQ8	Why do certain mutational events promote cancer phenotypes in some tissues and not in others?
9	driver mutations	2011	PQ9	As genomic sequencing methods continue to identify large numbers of novel cancer mutations, how can we identify the mutations in a given tumor that are most critical to the maintenance of its oncogenic phenotype?
10	epigenetic events	2011	PQ10	As we improve methods to identify epigenetic changes that occur during tumor development, can we develop approaches to discriminate between "driver" and "passenger" epigenetic events?
		2012	PQB2	
11	RNA processing	2011	PQ11	How do changes in RNA processing contribute to tumor development?
12	novel infectious agents	2011	PQ12	Given the recent discovery of the link between a polyomavirus and Merkel cell cancer, what other cancers are caused by novel infectious agents and what are the mechanisms of tumor induction?
13	improved in vivo imaging	2011	PQ13	Can tumors be detected when they are two to three orders of magnitude smaller than those currently detected with in vivo imaging modalities?
		2012	PQC5	
14	predicting progression	2011	PQ14	Are there definable properties of a non-malignant lesion that predict the likelihood of progression to invasive or metastatic disease?
		2012	PQC3	Are there definable properties of pre-malignant or other non-invasive lesions that predict the likelihood of progression to metastatic disease?
		2013	PQC1	What properties of pre-cancerous lesions or their microenvironment predict the likelihood of progression to malignant disease?
15	second primary cancers	2011	PQ15	Why do second, independent cancers occur at higher rates in patients who have survived a primary cancer than in a cancer-naïve population?
		2012	PQB1	
		2013	PQB1	
16	metastases clinical significance	2011	PQ16	How do we determine the clinical significance of finding cells from a primary tumor at another site?
		2012	PQC4	How do we determine the significance of finding cells from a primary tumor at another site and what methods can be developed to make this diagnosis clinically useful?

PQ	PRECIS	RFA YEAR	PQ ID IN RFA	QUESTION TEXT
17	new drug testing	2011	PQ17	Since current methods to assess potential cancer treatments are cumbersome, expensive, and often inaccurate, can we develop other methods to rapidly test interventions for cancer treatment or prevention?
		2012	PQD5	Since current methods to predict the efficacy or toxicity of new drug candidates in humans are often inaccurate, can we develop new methods to test potential therapeutic agents that yield better predictions of response?
18	undruggable targets	2011	PQ18	Are there new technologies to inhibit traditionally "undruggable" target molecules, such as transcription factors, that are required for the oncogenic phenotype?
19	chemotherapy sensitivity	2011	PQ19	Why are some disseminated cancers cured by chemotherapy alone?
		2012	PQD2	What molecular properties make some cancers curable with conventional chemotherapy?
		2013	PQD1	
20	immunotherapy biomarkers	2011	PQ20	Given the recent successes in cancer immunotherapy, can biomarkers or signatures be identified that can serve as predictors or surrogates of therapeutic efficacy?
21	therapy resistance	2011	PQ21	Given the appearance of resistance in response to cell killing therapies, can we extend survival by using approaches that keep tumors static?
		2012	PQD1	How does the selective pressure imposed by the use of different types and doses of targeted therapies modify the evolution of drug resistance?
22	oncogene addiction	2011	PQ22	Why do many cancer cells die when suddenly deprived of a protein encoded by an oncogene?
23	tumor indolence	2011	PQ23	Can we determine why some tumors evolve to aggressive malignancy after years of indolence?
		2012	PQC1	
24	metastasis study techniques	2011	PQ24	Given the difficulty of studying metastasis, can we develop new approaches, such as engineered tissue grafts, to investigate the biology of tumor spread?
		2012	PQB6	
25	physical activity & cancer	2012	PQA5	How does the level, type, or duration of physical activity influence cancer risk and prognosis?
		2013	PQA2	
26	susceptibility during development	2012	PQA6	How does susceptibility of exposure to cancer risk factors change during development?
		2013	PQA3	What biological mechanisms influence susceptibility to cancer risk factors at various stages of life?

PQ	PRECIS	RFA YEAR	PQ ID IN RFA	QUESTION TEXT
27	immune response	2012	PQB3	What molecular and cellular events determine whether the immune response to the earliest stages of malignant transformation leads to immune elimination or tumor promotion?
		2013	PQB2	What molecular and cellular events in the tumor microenvironment (for example, the local immune response) determine if a tumor at the earliest stages of malignant transformation is eliminated, stimulated for further development, or made indolent?
28	mutation or epigenetic change order	2012	PQB5	How does the order in which mutations or epigenetic changes occur alter cancer phenotypes or affect the efficacy of targeted therapies?
29	physical properties	2012	PQC2	How can the physical properties of tumors, such as a cell's electrical, optical or mechanical properties, be used to provide earlier or more reliable cancer detection, diagnosis, prognosis, or monitoring of drug response or tumor recurrence?
30	dormancy and recurrence	2012	PQC6	What molecular events establish tumor dormancy after treatment and what leads to recurrence?
		2013	PQC2	What molecular or cellular events establish tumor dormancy after treatment and what leads to recurrence?
31	long survivors	2012	PQD3	What underlying causal events—e.g., genetic, epigenetic, biologic, behavioral, or environmental—allow certain individuals to survive beyond the expected limits of otherwise highly lethal cancers?
32	cancer field effect	2012	PQD4	What properties of cells in a pre-malignant or pre-invasive field—sometimes described as the result of a cancer field effect—can be used to design treatments for a tumor that has emerged from this field or to block the appearance of future tumors?
		2013	PQA4	For tumors that arise from a pre-malignant field, what properties of cells in this field can be used to design strategies to inhibit the development of future tumors?
		2015	PQ1	
33	cachexia	2012	PQD6	What mechanisms initiate cachexia in cancer patients, and can we target them to extend lifespan and quality of life for cancer patients?
		2013	PQB3	What mechanisms initiate or sustain cancer cachexia, and can we target them to extend lifespan and quality of life for cancer patients?

APPENDIX B SUMMARY OF INTERVIEW / SURVEY PROCESS

With the help of an independent Evaluation Advisory Committee including members from the Science and Technology Policy Institute and other branches of NCI not involved in the PQ Initiative, the following evaluation areas were prioritized: PQ development process, application process, review process, management process, quality and scientific outcome of PQ research that include community enthusiasm, innovation and successes identified. The interview guide and surveys were developed by the Science and Technology Policy Institute and vetted by both the Evaluation Advisory Committee and a focus group in late 2015. The surveys and interview guides were cleared for use by the NCI Office of Management and Budget in February 2016. No statistical analysis was conducted on the interview and survey results.

Interviews

Nine members of the Provocative Questions Executive Committee were interviewed about the PQ development and retirement process from March 4, 2016 to April 12, 2016. The interviews lasted from 27 to 50 minutes.

Three Branch Chiefs and six Program Directors were interviewed about the PQ development process, portfolio administration, progress and outcomes. The interviews were conducted from March 8, 2016 through March 31, 2016, and lasted from 23 to 51 minutes.

Nine Workshop Participants were interviewed about the PQ process and outcomes from March 9, 2016 to April 21, 2016. The interviews lasted from 10 to 44 minutes. Five Interviewees participated in thematic workshops, three participated in regional workshops and two participated in the initial exploratory workshop (one Interviewee participated in both an exploratory and a thematic workshop).

Nine current or former Reviewers were interviewed about the grant review process from March 7, 2016 through April 15, 2016. The interviews lasted from 23 to 45 minutes.

Surveys

An online survey was conducted on PQ Applicants (i.e., those who applied but did not receive funding) and PQ Awardees, as identified from the NIH QVR system. The surveys sought responses on the Provocative Questions development process, the application process, and their perspective on the future of the PQ Initiative. Surveys were successfully delivered to 1,551 Applicants and 241 Awardees in the spring of 2016. A total of 720 (46.5% response rate) Applicants and 153 (63.1%) Awardees completed the survey.

APPENDIX C METHODS TO SELECT PUBLICATIONS RELEVANT TO EACH PQ

Since it would be impractical to review all scientific publications for inclusion in the evaluation, a mixture of machine learning and SME review techniques were employed to identify PQ-related publications. In this evaluation, an explicit assumption made is the number of PQ-related publications can be used as a proxy for the size of the research area. Machine learning techniques, and specifically in this evaluation, topic modeling techniques, provide a relatively fast way for human experts to explore and filter out irrelevant data. First, from the publication database, Cancer publications were identified for the period of interest. Then for each PQ, a broad set of candidate publications were identified. Within each candidate pool, topics were identified using a machine learning algorithm called Latent Dirichlet Allocation (LDA), and those topics that were deemed relevant to the PQ by SMEs were isolated. Lastly candidate publications that contain relevant topics were isolated as PQ-related publications. Publications selected using this method was used in analyses in Sections 3.1.2, 3.2.1, 3.2.2, 3.3.1, 3.3.2, Appendix D, Appendix F and Appendix G. Below is a detailed description of the steps involved.

Selection of candidate papers by Web of Science Subject Categories and keywords

Cancer publications were identified using an assortment of keywords as defined and approved by the NCI Provocative Questions evaluation team in the previous evaluation.

Cancer publications as indexed in the Web of Science database were identified following 3 rules as defined and approved by the NCI Provocative Questions evaluation team in the previous evaluation: (1) the publication FY must be 2008 or later, (2) the article must be in journal associated with at least 1 of the 55 subject categories listed in the table below,[24] (3) the article's abstract text must contain at least 2 out of the 47 general cancer terms, with an asterisk (*) indicating a wildcard match: angiosarcoma , anti*cancer* , anti*neoplas* , anticancer* , antineoplas* , blastoma* , cancer* , carcino* , cell* transform* , chondrosarcoma , cystosarcoma , dermatofibrosarcoma , endothelioma , epithelioma , fibrosarcoma , glioma , hemangiopericytoma , hemangiosarcoma , hepatoblastoma , histiocytoma , histiocytoma* , leiomyosarcoma , leukaemi* , leukemi* , liposarcoma , lymphangiosarcoma , lymphoma , lymphosarcoma , malignan* , medulloblastoma , melanom* , metastati* , myeloma , neoplas* , nephroblastoma , neuroblastoma , neurofibrosarcoma , oncogen* , oncolog* , osteosarcoma , pancreatoblastoma , pleuropulmonary blastoma , retinoblastoma , rhabdomyosarcoma , sarcom* , tumor* , tumour*.

INCLUDED WEB OF SCIENCE SUBJECT CATEGORIES	
Behavioral Sciences	Medical Laboratory Technology
Biochemical Research Methods	Medicine, General & Internal
Biochemistry & Molecular Biology	Medicine, Miscellaneous
Biology	Medicine, Research & Experimental
Biology, Miscellaneous	Microbiology
Biophysics	Microscopy
Biotechnology & Applied Microbiology	Multidisciplinary Sciences

[24] The Web of Science scheme comprises more than 250 subject areas in science, social sciences, and arts & humanities. Many broad areas such as physics and materials science are represented by smaller subfields.

INCLUDED WEB OF SCIENCE SUBJECT CATEGORIES

Cell & Tissue Engineering	Nanoscience & Nanotechnology
Cell Biology	Nutrition & Dietetics
Chemistry, Analytical	Oncology
Chemistry, Applied	Parasitology
Chemistry, Inorganic & Nuclear	Pediatrics
Chemistry, Medicinal	Pharmacology & Pharmacy
Chemistry, Multidisciplinary	Physiology
Chemistry, Organic	Psychology, Applied
Chemistry, Physical	Psychology, Biological
Clinical Neurology	Psychology, Clinical
Cytology & Histology	Psychology, Developmental
Developmental Biology	Psychology, Educational
Engineering, Biomedical	Psychology, Experimental
Evolutionary Biology	Psychology, Multidisciplinary
Genetics & Heredity	Psychology, Social
Health Care Sciences & Services	Radiology, Nuclear Medicine & Medical Imaging
Health Policy & Services	Social Sciences, Biomedical
Integrative & Complementary Medicine	Spectroscopy
Materials Science, Biomaterials	Toxicology
Mathematical & Computational Biology	Virology
Medical Informatics	

Cancer publications were further stratified into candidate publications relevant to each specific PQ from 2011, 2012, and 2013 issuances (33 total). For each PQ, a set of PQ-specific keywords was defined by the NCI Provocative Questions evaluation team based on the text of the RFA of this PQ. Using these keywords, candidate publications of each PQ were identified. In addition, funded publications and Gold Standard publications that were identified previously were added to each PQ's candidate corpus if not already included via keyword matching.

Identification of topics covered by candidate papers using topic models

LDA[25], a widely-used machine learning algorithm was used to identify, preliminarily, potential topics represented by the candidate publication in each PQ. Clarivate Analytics set up software to process candidate publications metadata (article title, abstract, author-submitted keywords, and MeSH terms), eliminate common words, and identify 100 topics for each PQ from the remainder text[26]. The choice of 100 topics was determined by trial-and-error to provide sufficient granularity without unnecessary division of a coherent concept in a topic. A

[25] Blei D, Ng A, Jordan M. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3:993–1022
[26] The statistical software R (ver. 3.2.1) with packages "lda" (ver. 1.3.2) and "LDAvis" (ver. 0.3.2) were used.

model with over 90% of coherent topics was considered as a good model. An exception of 50 topics was chosen for PQ 18, however, because there were only 385 candidate publications identified.

SMEs' review of machine identified topics for each PQ

To determine the quality of topic models and to identify the topics that may be related to PQs, Clarivate Analytics conducted preliminary evaluations on the topic model output to eliminate incoherent topics (e.g., topics whose keywords cannot easily be interpreted to describe a single concept), and to determine which topics were relevant to specific PQs. The preliminary results produced by Clarivate analysts were then subjected to evaluation by NCI SMEs. NCI SMEs identified topics related to their respective expertise areas of study. When possible, SMEs were asked to quantify the relatedness of topics to a PQ or to select the most relevant topics. The decision and feedback provided by the SMEs for each PQ constituted the topic inclusion rules to identify PQ-related publications, described below.

Selection of PQ-related publication by topic inclusion rules

The LDA topic modeling method assigns a probability for each of the 100 topics, or 50 topics in case of PQ18, to the paper. This means that each candidate paper of a PQ is assigned to a mixture of multiple topics. Since 100 topics were typically generated from a pool of many thousands of papers, a single paper may not contain all topics. In fact, a paper may contain approximately 15 topics, most of which are associated with the paper with low probabilities.

LDA generated topics were reviewed for their relatedness to each PQ at the level of each PQ. Due to each PQ's unique nuances, these topics may be standalone or combined as suggested by the SMEs in order to select PQ-related papers.[27] There are three different cases to select PQ-related papers.

- Case 1. The subject matter of a PQ contains several, more or less independent sub-concepts. No single topic covers all sub-concepts, but there are multiple topics each of which corresponds to a sub-concept.

  Example:

  PQ 11. How do changes in RNA processing contribute to tumor development?

  In this PQ, the subject of RNA processing includes several sub-concepts such as "Alternative RNA splicing," "RNA Interference," "miRNA and lncRNA processing," etc. Each of these sub-concepts was identified as a PQ-related topic. Papers that contain at least one such topic were selected as PQ-related papers.

  Nineteen PQs belong to case 1. They are PQs 1, 2, 3, 4, 5, 6, 11, 12, 13, 16, 17, 18, 22, 23, 25, 29, 30, 32, and 33.

- Case 2. The subject matter that a PQ intends to study is a research area with specified scope. Again, no single topic covers the nuance of the PQ well, but there are topics that are related to the research area and topics that are related to the scope. In this case, papers that contain both topics were selected as PQ-related topics.

  Example:

[27] A candidate paper that bears a non-zero probability to a PQ-related topic may be related to the PQ. However, if the probability of association with a PQ-related topic is low, the relatedness may be trivial. Therefore, we only considered a paper's top 3 topics by probabilities of association to determine whether it was related to the PQ.

PQ 20. Given the recent successes in cancer immunotherapy, can biomarkers or signatures be identified that can serve as predictors or surrogates of therapeutic efficacy?

This PQ is intended to study biomarkers, not all biomarkers, but those biomarkers that have utilities in predicting effectiveness of cancer immunotherapy. The SME had identified a group of topics that related to cancer immunotherapy, such as "immune checker blockade mAb," "proinflammatory cytokines," "Tumor infiltrating lymphocytes" and "Cancer vaccines." None of these topics contains information about biomarkers. Therefore, only papers containing at least one cancer immunotherapy topic as well as one of biomarker topic, such as "Tumor biomarkers," "Tumor associated antigens," and "Immunoassay for cancer immunotherapy" were considered as PQ-related papers.

Eleven PQs belong to case 2. They are PQs 7, 8, 9, 14, 15, 19, 20, 21, 24, 26, and 31.

- Case 3. Papers that are so nuanced that none of the topic or topic combinations exactly matched to the PQ

Example:

PQ 28. How does the order in which mutations or epigenetic changes occur alter cancer phenotypes or affect the efficacy of targeted therapies?

This PQ is intended to study not just the mutation or epigenetic changes, but also the order of these events. Analysis by the Clarivate analyst and SME only identified topics that are related to the mutations or epigenetic changes which affect cancer phenotypes or the efficacy of targeted therapies. But, none of the top 30 most common keywords of these topics is about the orders of these molecular events. The review of the funded papers suggested that order information is implicit and less likely to be detected by LDA topic modeling. In these cases, we requested the SMEs to identify the topics that are most related to the topics. From these selected topics we include publications as PQ-related publications.

Other PQs in this category are PQ 10 and 27.

Before and After PQ designation

PQ-related publications were divided by the year of their publications. For each PQ, those publications that were published prior to and in the first year that PQ was issued were considered to have been published "before PQ." For those that were published after the year in which PQ was issued were considered "after PQ."

APPENDIX D FRACTIONAL SHARE OF CANCER PUBLICATIONS IN EACH PQ'S RESEARCH AREA BEFORE AND AFTER THE ISSUANCE OF EACH QUESTION

The overall volume of Cancer literature has grown linearly over the years. We computed publication share to attempt to account for the baseline increase of publications. In addition, the change in the share of Cancer research authors active in a given PQ research area is provided. Distinct authors' names were compiled across all PQ-related publications and counted for each PQ to estimate the number of active authors in each PQ targeted research area. It should be noted that a full disambiguation of the author names was not performed. For example, "Robert Johnson" and "Rob Johnson" may count as two authors while in actuality they are the same person. Thus the absolute number of authors identified here is likely to be an over-estimate. The direction of change, however, should remain the same.

PQ	SHARE OF CANCER PUBLICATIONS		PERCENT CHANGE IN PUBLICATION SHARE	PUBLICATION SHARE CHANGE P-VALUE[28]		PERCENT CHANGE IN SHARE OF CANCER RESEARCH AUTHORS[29]
	Before PQ	After PQ				
1 - obesity and cancer	0.11%	0.15%	33.68%	<0.001	**	41.77%
2 - geographic environmental risks	0.09%	0.10%	3.95%	0.5756		11.30%
3 - measuring risk exposure	0.67%	0.69%	2.41%	0.35821		6.83%
4 - cognitive processes for behavior change	0.04%	0.05%	27.83%	0.01253	*	78.21%
5 - drugs for other indications	0.36%	0.44%	21.07%	<0.001	**	51.77%
6 - disease correlation	0.20%	0.22%	7.03%	0.14402		23.04%
7 - age dependence	0.54%	0.35%	-34.35%	<0.001	**	-23.57%
8 - tissue-dependent phenotypes	0.01%	0.00%	-23.91%	0.08682		-9.33%
9 - driver mutations	0.07%	0.04%	-45.97%	<0.001	**	-37.31%
10 - epigenetic events	1.30%	1.17%	-10.10%	<0.001	**	0.10%
11 - RNA processing	2.02%	2.30%	13.77%	<0.001	**	11.15%
12 - novel infectious agents	0.34%	0.40%	16.81%	<0.001	**	34.35%
13 - improved in vivo imaging	0.08%	0.08%	-1.84%	0.8065		6.63%
14 - predicting progression	0.37%	0.41%	12.12%	<0.001	**	15.34%
15 - second primary cancers	0.12%	0.09%	-28.09%	<0.001	**	-10.99%

[28] A likelihood ratio test was used to determine the statistical significance of percentage change.

PQ	SHARE OF CANCER PUBLICATIONS		PERCENT CHANGE IN PUBLICATION SHARE	PUBLICATION SHARE CHANGE P-VALUE[28]		PERCENT CHANGE IN SHARE OF CANCER RESEARCH
16 - metastases clinical significance	1.63%	2.38%	46.26%	<0.001	**	57.81%
17 - new drug testing	0.51%	0.60%	17.00%	<0.001	**	22.58%
18 - undruggable targets	0.04%	0.03%	-6.62%	0.54862		8.63%
19 - chemotherapy sensitivity	0.04%	0.06%	31.69%	0.00324	**	36.81%
20 - immunotherapy biomarkers	0.12%	0.09%	-20.31%	<0.001	**	-20.79%
21 - therapy resistance	0.03%	0.10%	258.33%	<0.001	**	395.22%
22 - oncogene addiction	0.42%	0.33%	-21.37%	<0.001	**	-10.18%
23 - tumor indolence	0.14%	0.16%	9.37%	0.1048		21.53%
24 - metastasis study techniques	0.19%	0.24%	28.71%	<0.001	**	37.16%
25 - physical activity & cancer	0.32%	0.38%	18.96%	<0.001	**	16.37%
26 - susceptibility during development	0.15%	0.10%	-30.46%	<0.001	**	-24.76%
27 - immune response	0.01%	0.01%	-13.17%	0.69395		-6.38%
28 - mutation or epigenetic change order	0.01%	0.02%	70.90%	0.00399	**	42.78%
29 - physical properties	0.26%	0.27%	3.36%	0.48449		-8.14%
30 - dormancy and recurrence	0.18%	0.20%	15.43%	0.00937	**	14.83%
31 - long survivors	0.01%	0.01%	36.20%	0.15158		141.55%
32 - cancer field effect	0.19%	0.13%	-29.45%	<0.001	**	-25.73%
33 – cachexia	0.02%	0.04%	117.92%	<0.001	**	150.23%

APPENDIX E PREDICTIVE ESTIMATES OF PQ PROGRAM, ACTIVITY CODE, FY AND CRITERION SCORE ON OVERALL IMPACT SCORE

A linear regression model, estimating the contribution to the overall impact score by the following factors: the five criterion scores, whether the application applied to a PQ RFA, activity code (R01 or R21) of the applications, the FY of applications, the interaction effect between application year and criterion scores, and the interaction between activity code and criterion scores. The regression was performed using R with the stats package (ver. 3.2.1).

The result is presented below. The estimate results revealed there a reduction of the Approach score (better) would reduce the impact score by 7.34 when all else being equal. However, there is a strongly significant interaction effect from Approach and PQ application ("Approach:isPQ"). That is, a reduction of the Approach score (better) would reduce the impact score by 2.22 less (i.e., 7.34-2.22=5.12) for a PQ application than a typical application when all else being equal.

FACTORS	ESTIMATE	P-VAL	SIG
(Intercept)	-3.89427	< 0.001	***
Approach	7.33917	< 0.001	***
Significance	2.8882	< 0.001	***
Innovation	1.96839	< 0.001	***
Investigator	0.3952	<0.01	**
Environment	-0.19155		
is PQ	12.07021	< 0.001	***
is R21	-1.37912	<0.01	**
FY 2012	0.4968	< 0.001	***
FY 2013	0.8444	< 0.001	***
FY 2014	0.68863	< 0.001	***

FACTORS	ESTIMATE	P-VAL	SIG
Approach:isPQ	-2.2209	< 0.001	***
Significance:isPQ	-0.75886	<0.05	*
Innovation:isPQ	-0.49485	<0.1	
Investigator:isPQ	-0.56275		
Environment:isPQ	0.3153		
Approach:isR21	-0.41768	<0.01	**
Significance:isR21	0.06036		
Innovation:isR21	0.32626	<0.05	*
Investigator:isR21	0.68605	< 0.001	***
Environment:isR21	-0.2381		

Residual standard error: 6.317 on 17355 degrees of freedom

R^2: 0.72

F-statistic: 2279 on 20 and 17355 DF, p-value: < 2.2e-16

APPENDIX F PERCENT OF PIS NEW TO NCI IN EACH PROVOCATIVE QUESTION DURING THE STUDY PERIOD

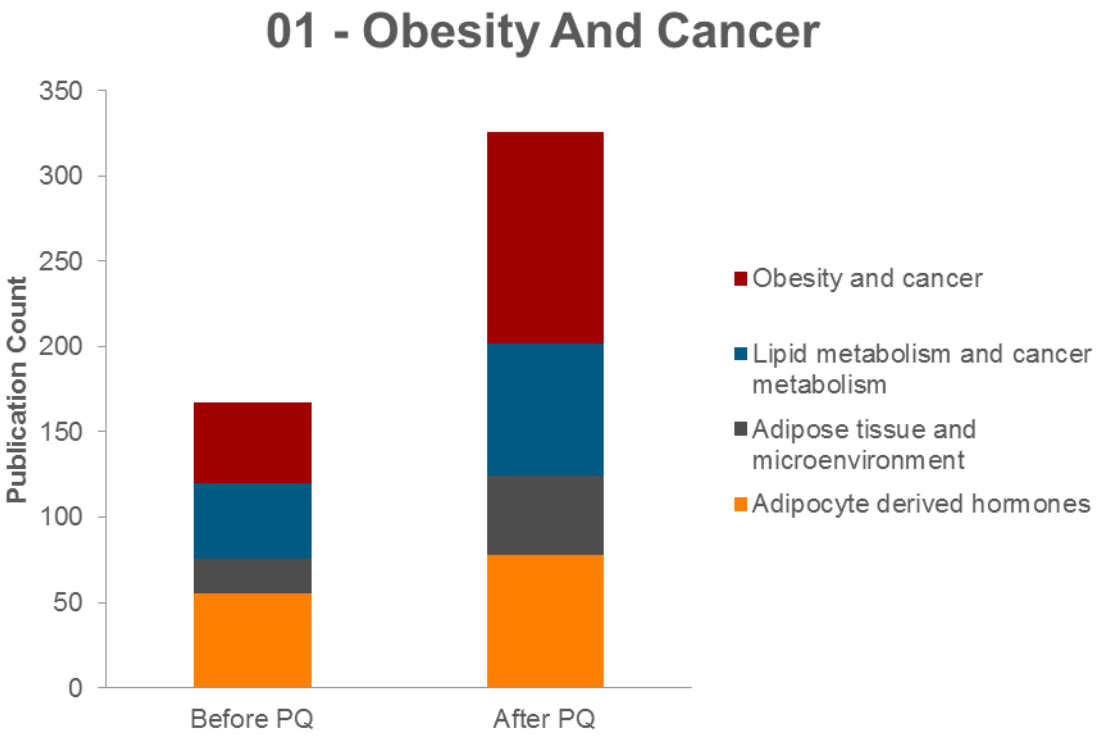
The author names of PQ-related publications were crossed referenced with PI names of those who received NCI research funding between 2001 and 2015, as reported in NIHRePORTER. This gave us the total PI count. Those who did not receive NCI funding until the first PQ issuance were then counted as PIs new to NCI. Thus an investigator who had received funding from an Institute other than NCI but only received NCI funding after PQ issuance would still be considered someone new to NCI. Similar statistics were also computed for NCI as a whole. The table below lists PQs in the order of the percentage of new PIs to NCI.

PQ	NEW PI COUNT	TOTAL PI COUNT	PERCENT OF PIS NEW TO NCI
8 - tissue-dependent phenotypes	2	5	40.00%
27 - immune response	2	8	25.00%
28 - mutation or epigenetic change order	6	26	23.08%
21 - therapy resistance	11	52	21.15%
18 - undruggable targets	8	42	19.05%
17 - new drug testing	77	453	17.00%
16 - metastases clinical significance	136	917	14.83%
NCI	*1,562*	10,857	*14.39%*
14 - predicting progression	40	285	14.04%
24 - metastasis study techniques	29	207	14.01%
20 - immunotherapy biomarkers	8	59	13.56%
23 - tumor indolence	18	133	13.53%
29 - physical properties	18	138	13.04%
33 – cachexia	2	16	12.50%
4 - cognitive processes for behavior change	3	25	12.00%
3 - measuring risk exposure	30	258	11.63%
19 - chemotherapy sensitivity	3	27	11.11%
30 - dormancy and recurrence	18	163	11.04%
22 - oncogene addiction	40	364	10.99%
12 - novel infectious agents	22	220	10.00%
10 - epigenetic events	59	604	9.77%
5 - drugs for other indications	21	232	9.05%
11 - RNA processing	65	729	8.92%
1 - obesity and cancer	7	79	8.86%
13 - improved in vivo imaging	4	47	8.51%
9 - driver mutations	4	50	8.00%
32 - cancer field effect	12	156	7.69%
6 - disease correlation	4	58	6.90%
2 - geographic environmental risks	2	29	6.90%
7 - age dependence	21	307	6.84%
26 - susceptibility during development	4	59	6.78%
25 - physical activity & cancer	9	182	4.95%

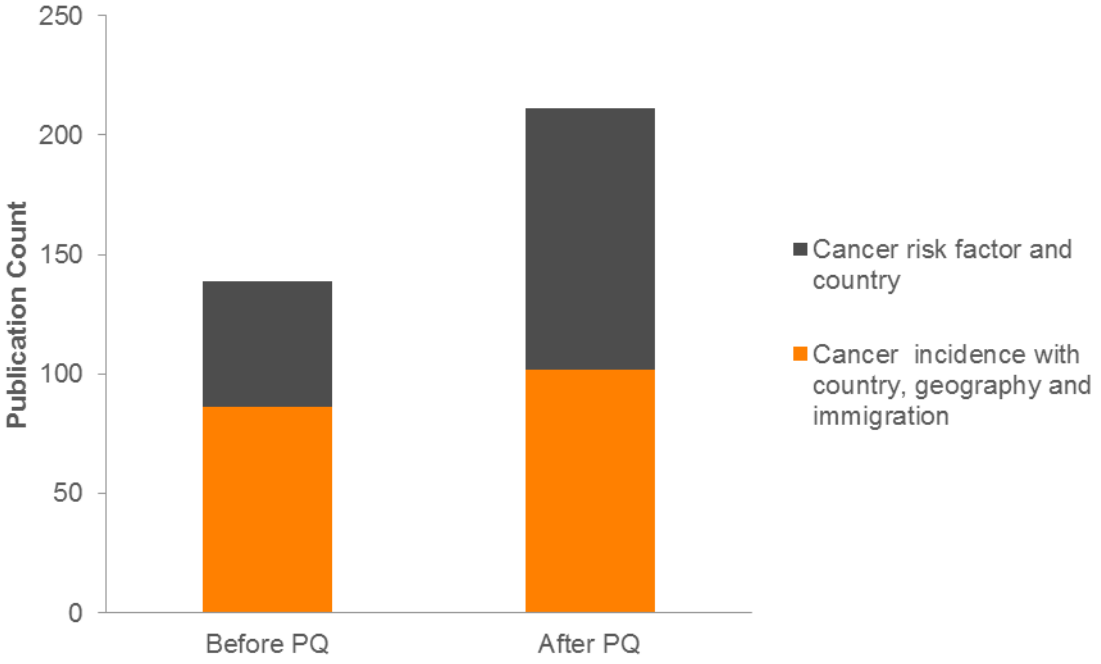
PQ	NEW PI COUNT	TOTAL PI COUNT	PERCENT OF PIS NEW TO NCI
15 - second primary cancers	2	61	3.28%
31 - long survivors	0	0	0.00%

APPENDIX G PUBLICATION COUNTS FOR EACH PQ

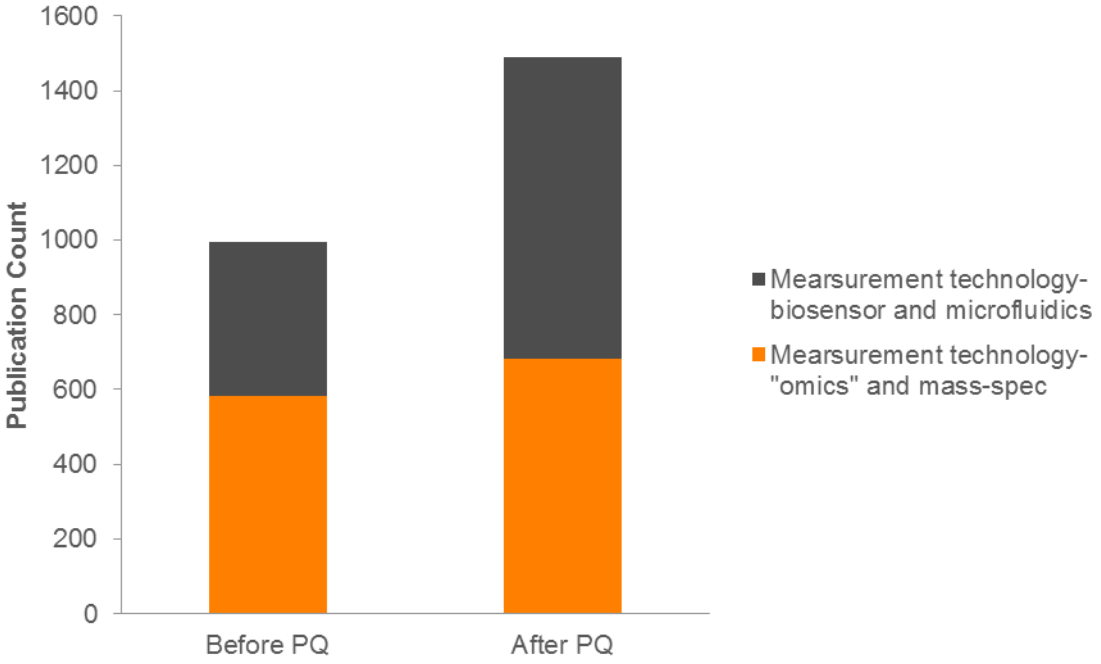
Fractional publication count by topic(s) for each PQ is shown below. If a publication covers *N* topics, it is counted *1/N* times in each topic. The method used to assign topics can be found in Appendix C.

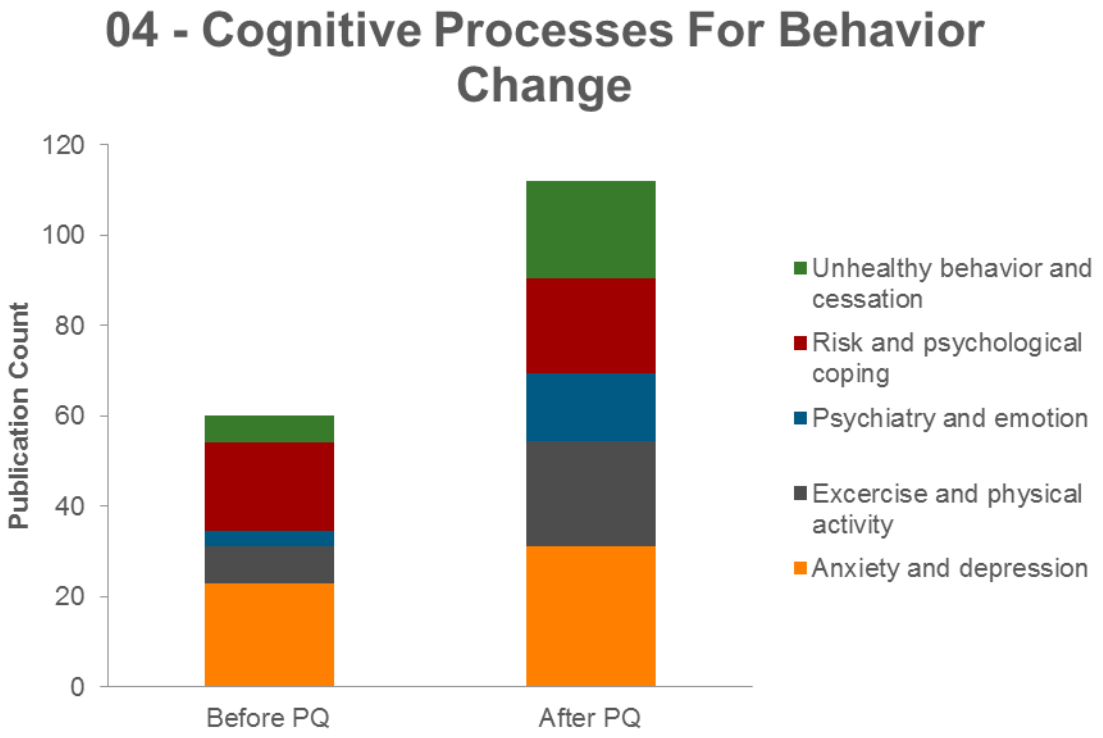


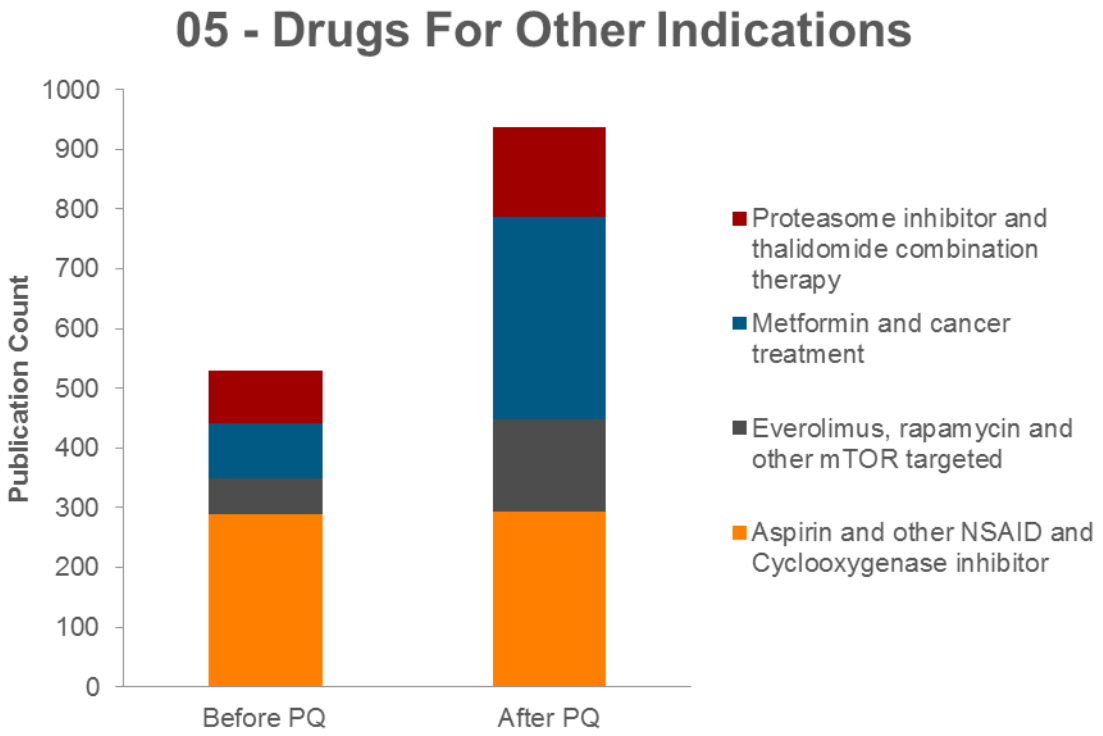
02 - Geographic Environmental Risks



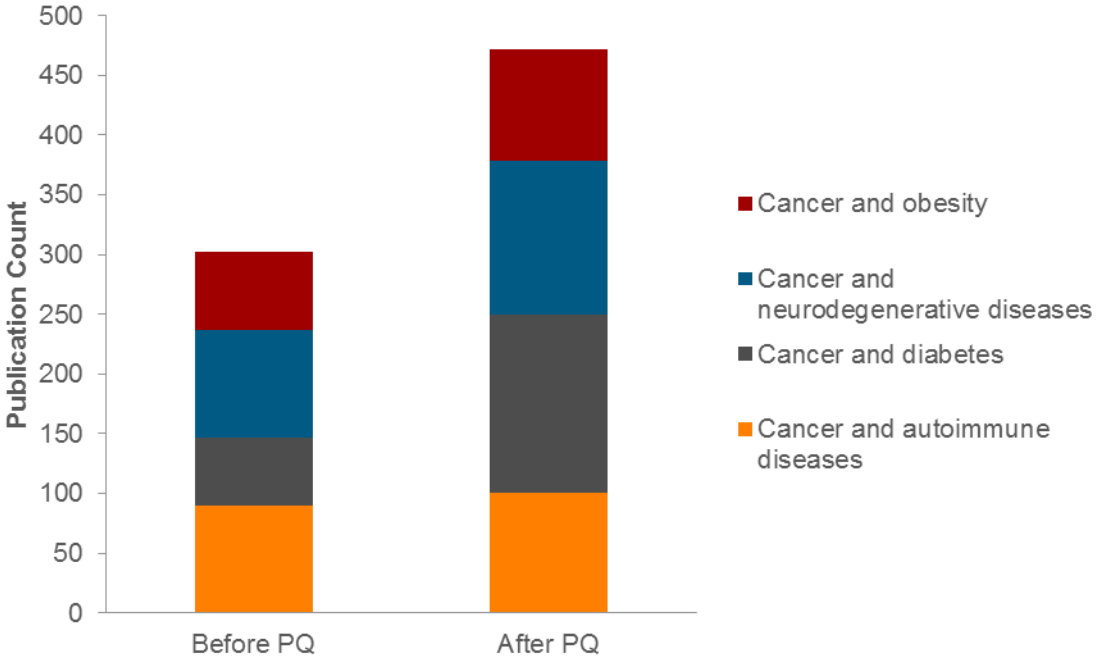
03 - Measuring Risk Exposure



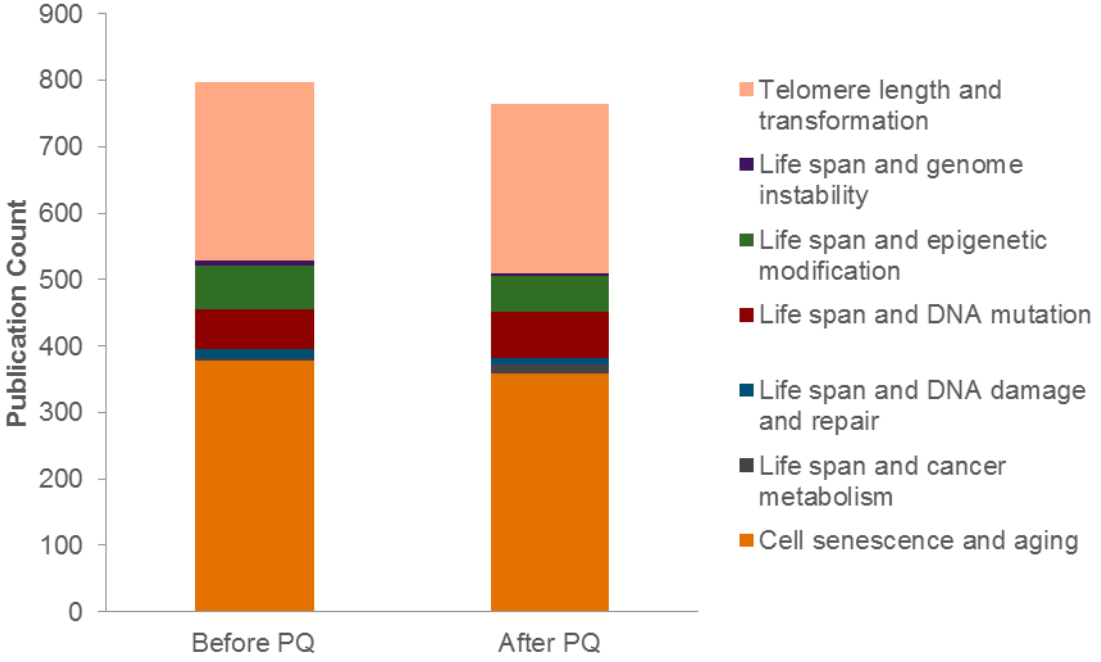

04 - Cognitive Processes For Behavior Change
Publication Count
120
100
80
60
40
20
0
Before PQ
After PQ
Unhealthy behavior and cessation
Risk and psychological coping
Psychiatry and emotion
Excercise and physical activity
Anxiety and depression

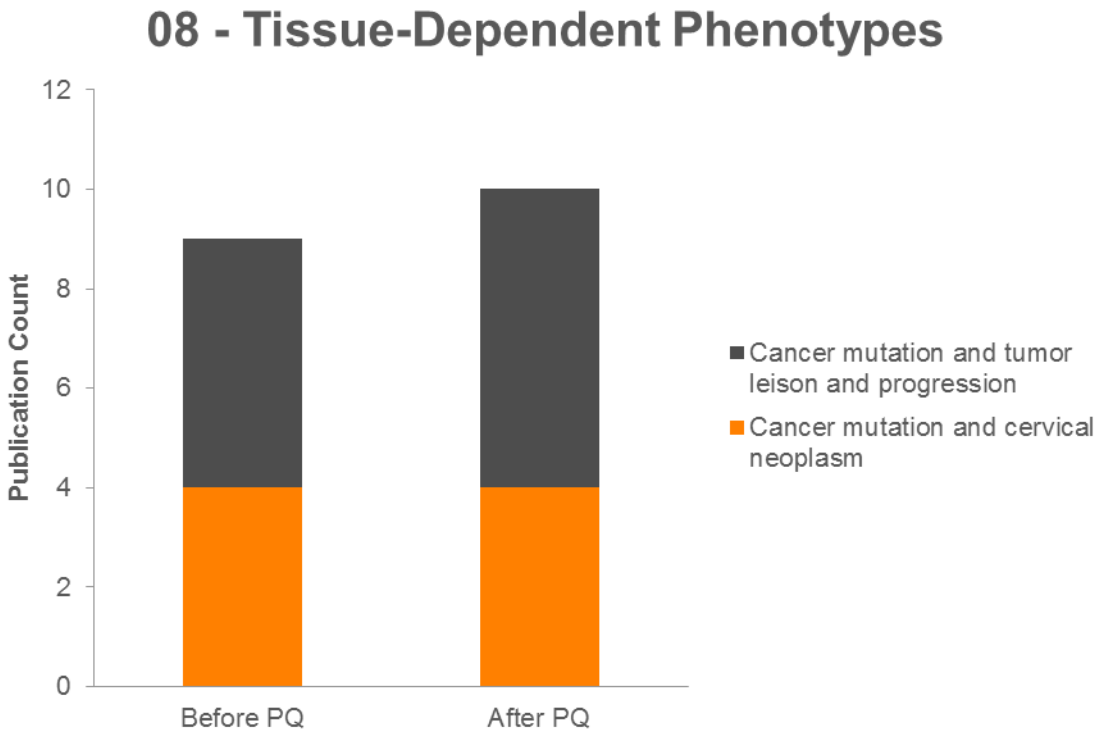

05 - Drugs For Other Indications
Publication Count
1000
900
800
700
600
500
400
300
200
100
0
Before PQ
After PQ
Proteasome inhibitor and thalidomide combination therapy
Metformin and cancer treatment
Everolimus, rapamycin and other mTOR targeted
Aspirin and other NSAID and Cyclooxygenase inhibitor

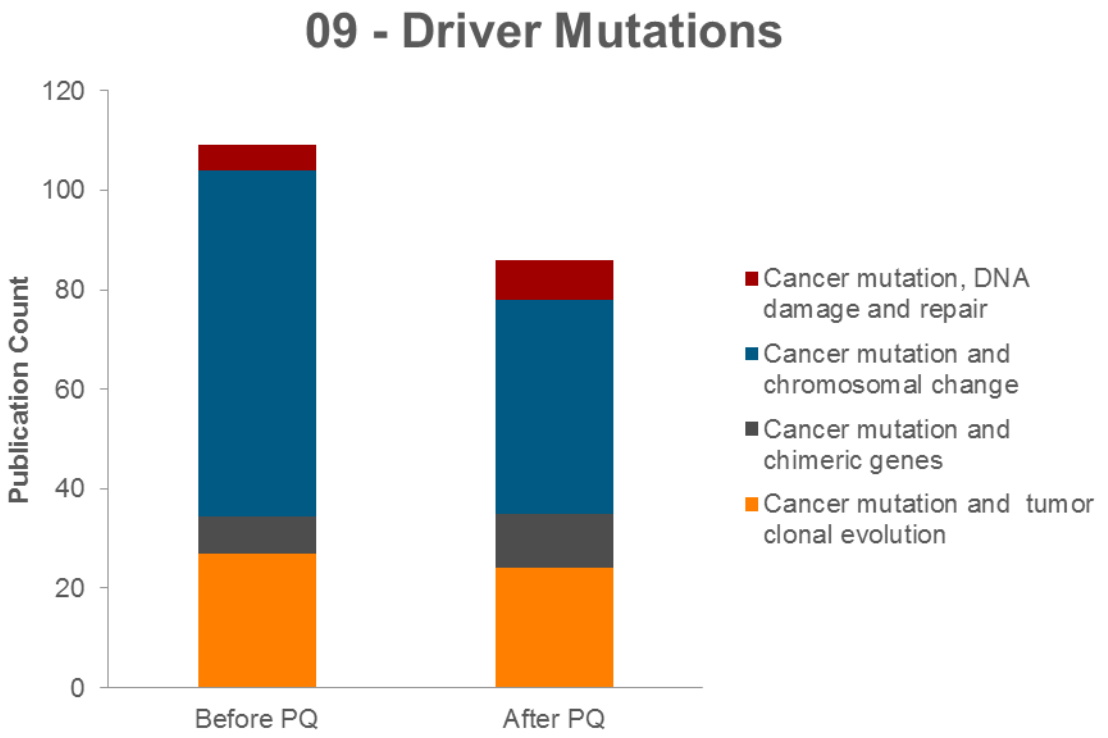
06 - Disease Correlation

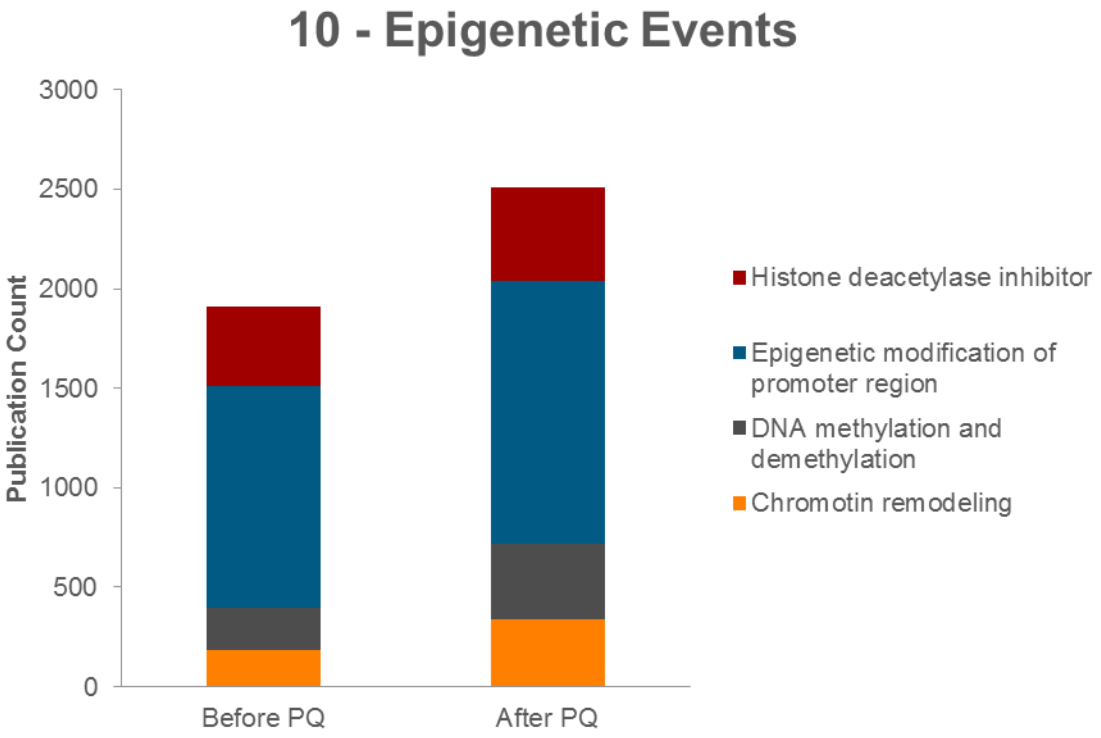


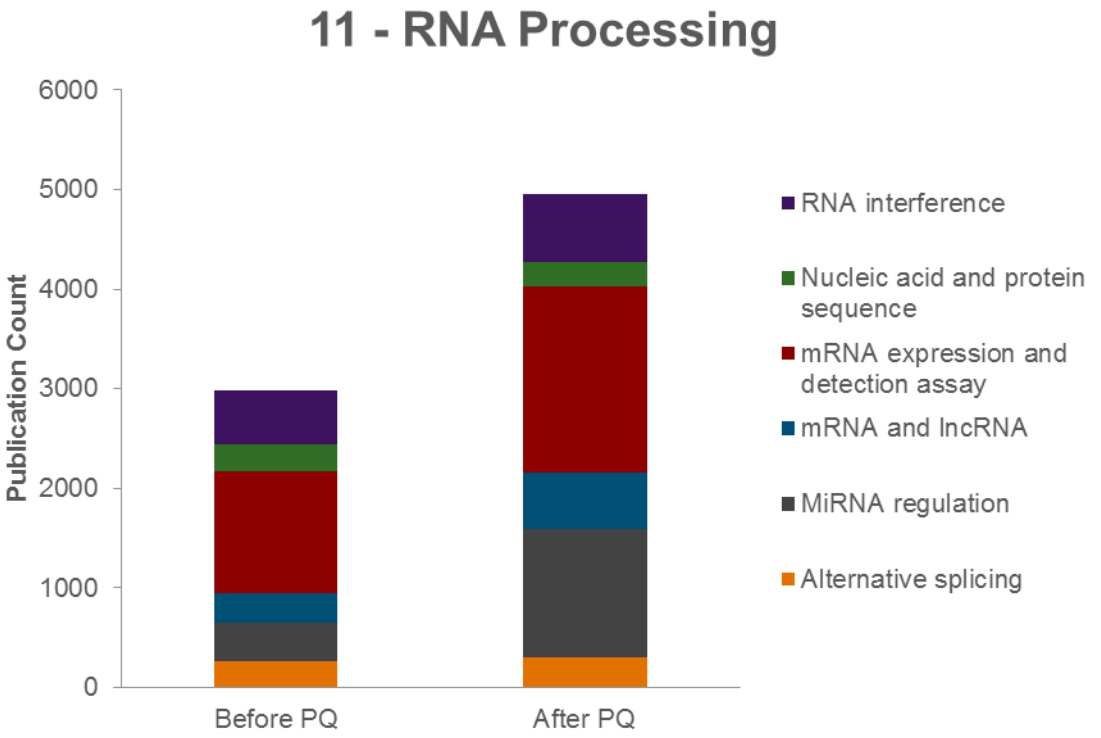
07 - Age Dependence



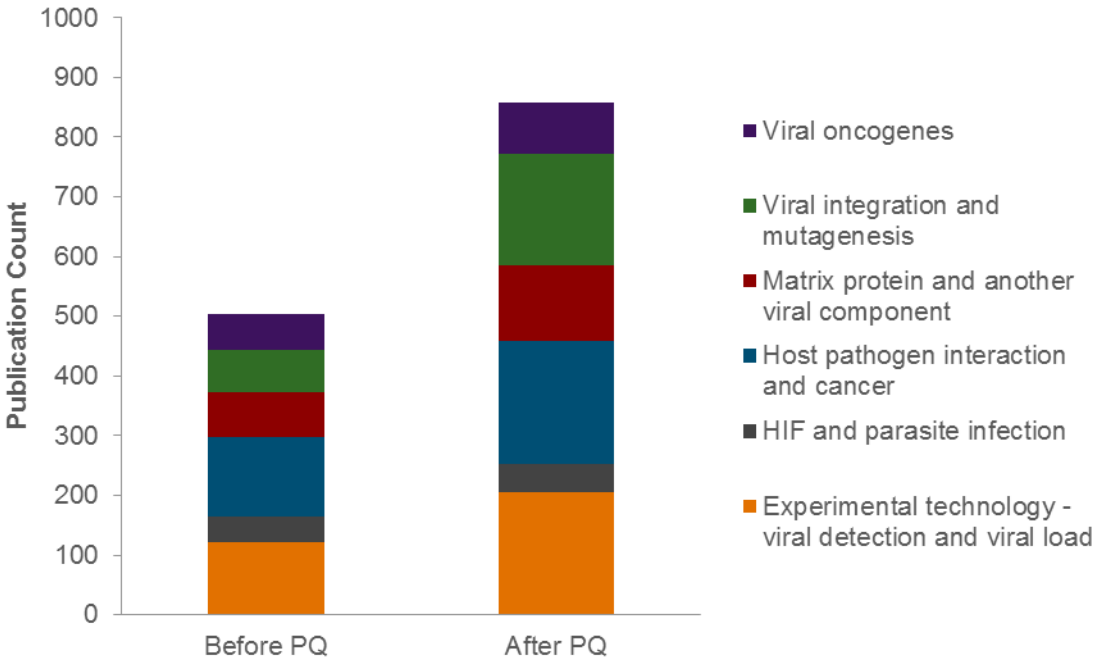

08 - Tissue-Dependent Phenotypes
Publication Count
12
10
8
6
4
2
0
Before PQ
After PQ
Cancer mutation and tumor leison and progression
Cancer mutation and cervical neoplasm


09 - Driver Mutations
Publication Count
120
100
80
60
40
20
0
Before PQ
After PQ
Cancer mutation, DNA damage and repair
Cancer mutation and chromosomal change
Cancer mutation and chimeric genes
Cancer mutation and tumor clonal evolution

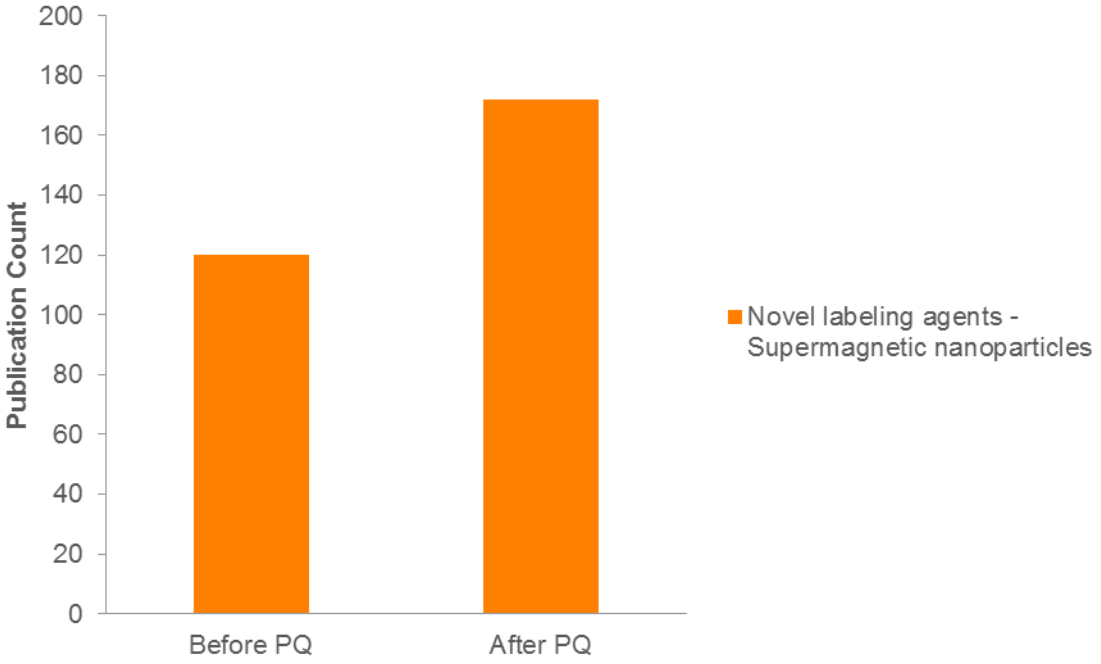

10 - Epigenetic Events
Publication Count
3000
2500
2000
1500
1000
500
0
Before PQ
After PQ
Histone deacetylase inhibitor
Epigenetic modification of promoter region
DNA methylation and demethylation
Chromotin remodeling

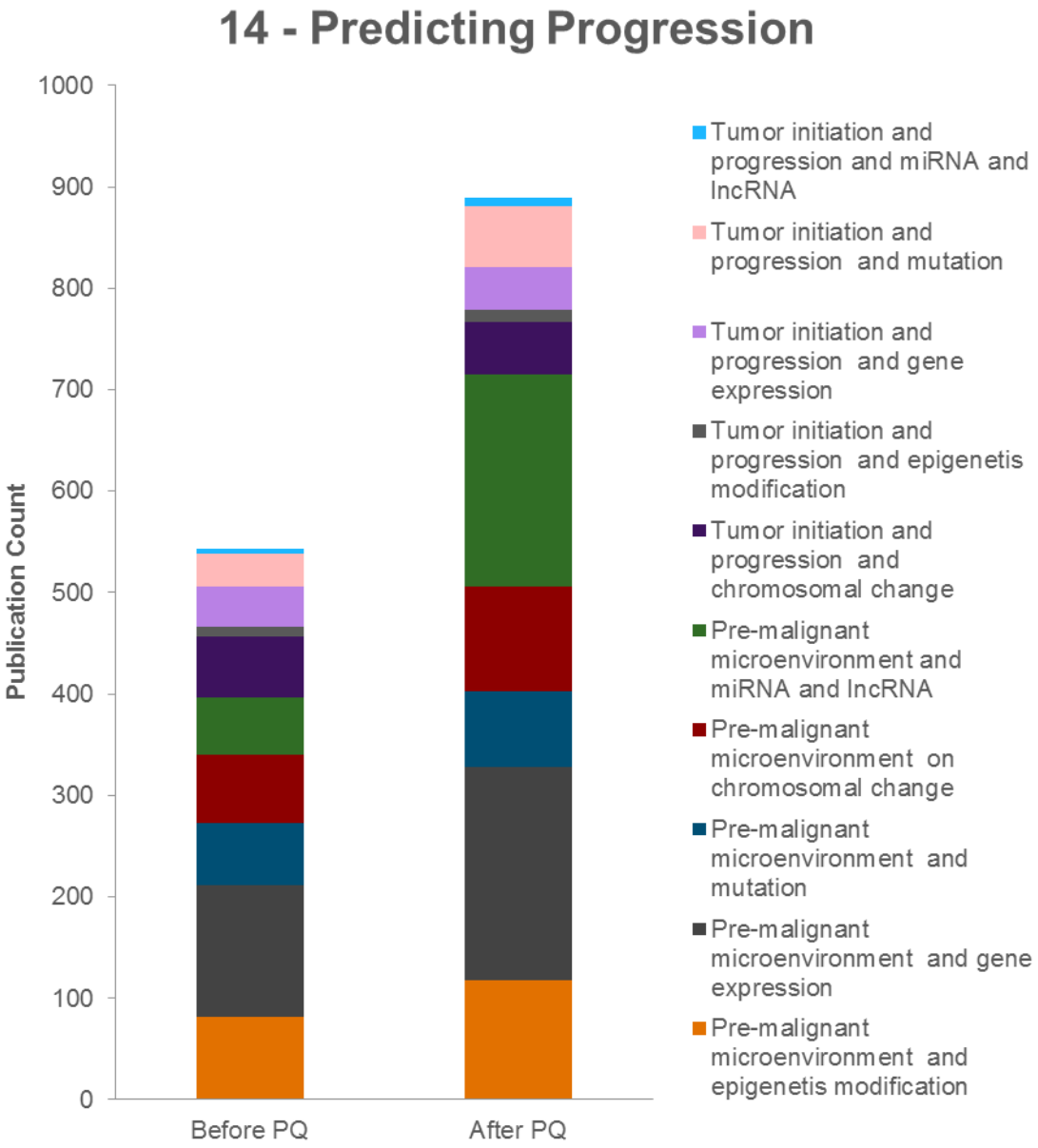

11 - RNA Processing
Publication Count
6000
5000
4000
3000
2000
1000
0
Before PQ
After PQ
RNA interference
Nucleic acid and protein sequence
mRNA expression and detection assay
mRNA and lncRNA
MiRNA regulation
Alternative splicing

12 - Novel Infectious Agents

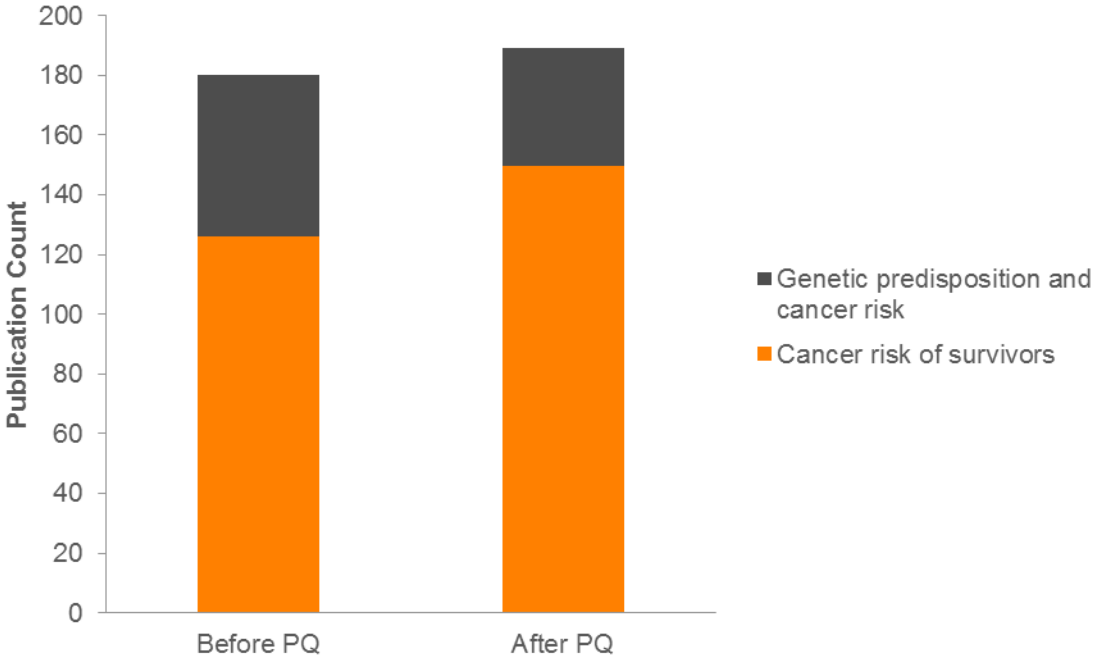


13 - Improved In Vivo Imaging

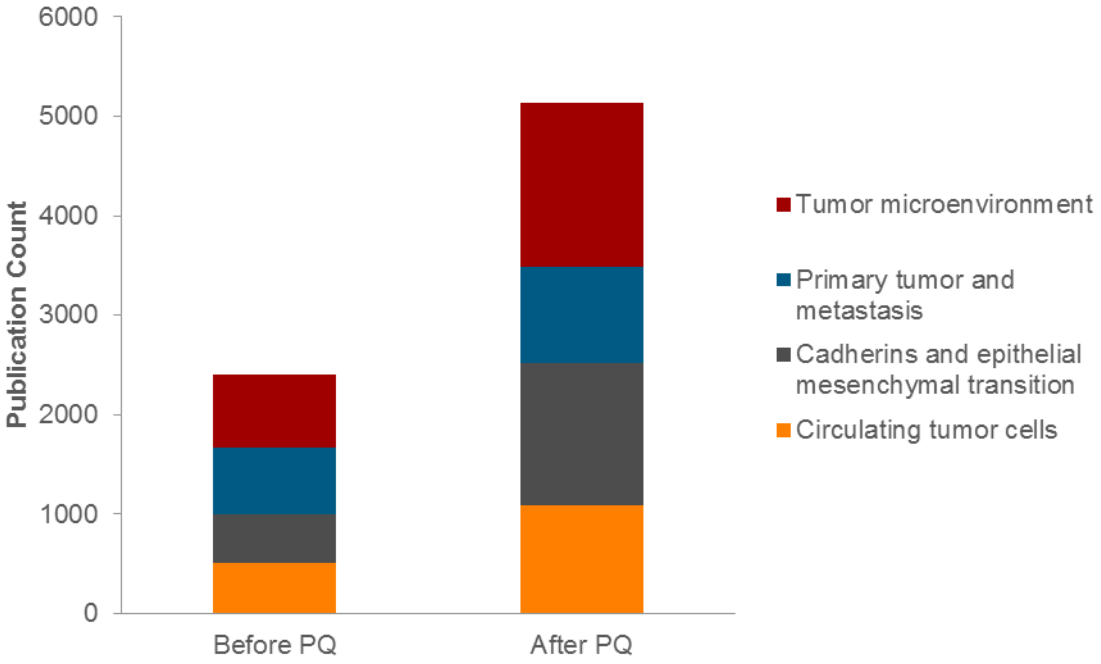



14 - Predicting Progression
Publication Count
1000
900
800
700
600
500
400
300
200
100
0
Before PQ
After PQ
Tumor initiation and progression and miRNA and lncRNA
Tumor initiation and progression and mutation
Tumor initiation and progression and gene expression
Tumor initiation and progression and epigenetis modification
Tumor initiation and progression and chromosomal change
Pre-malignant microenvironment and miRNA and lncRNA
Pre-malignant microenvironment on chromosomal change
Pre-malignant microenvironment and mutation
Pre-malignant microenvironment and gene expression
Pre-malignant microenvironment and epigenetis modification

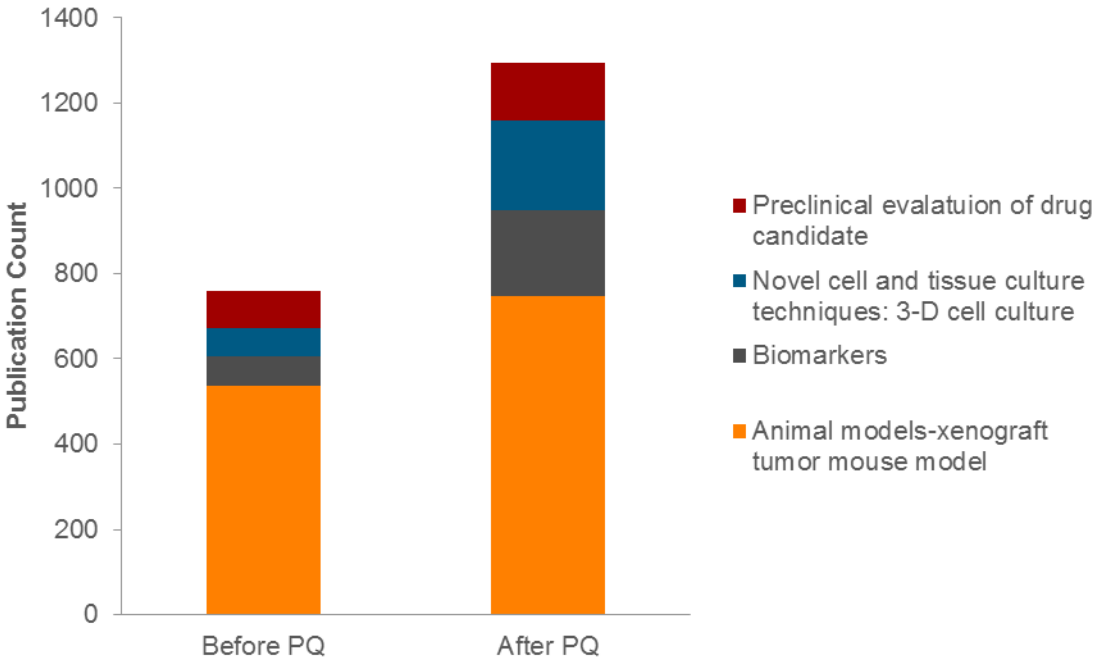
15 - Second Primary Cancers



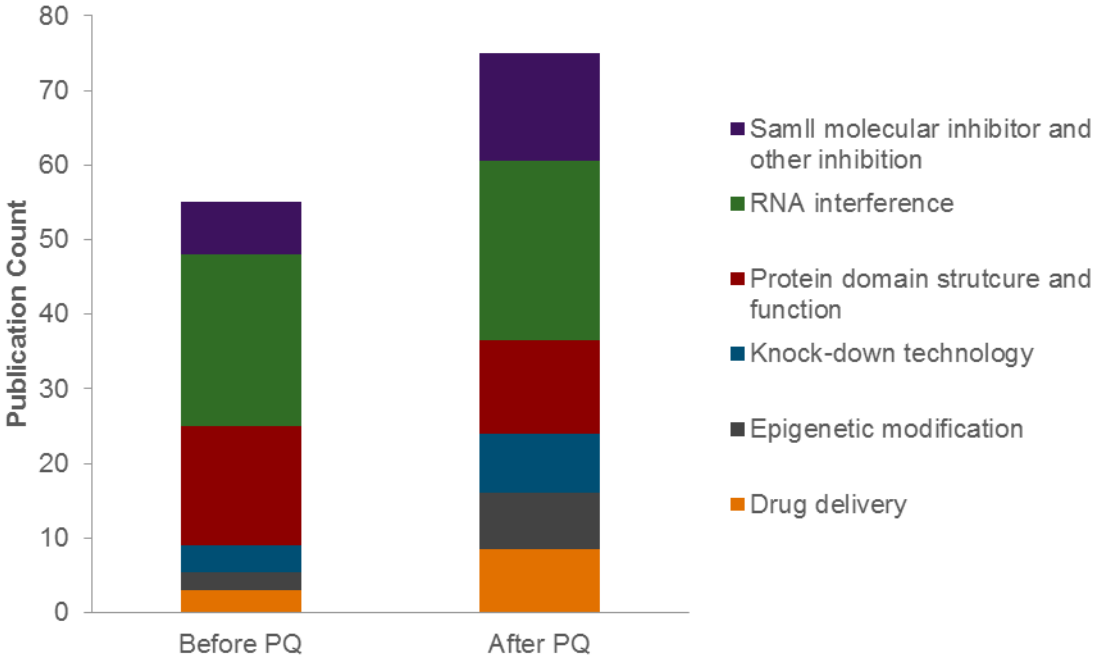
16 - Metastases Clinical Significance

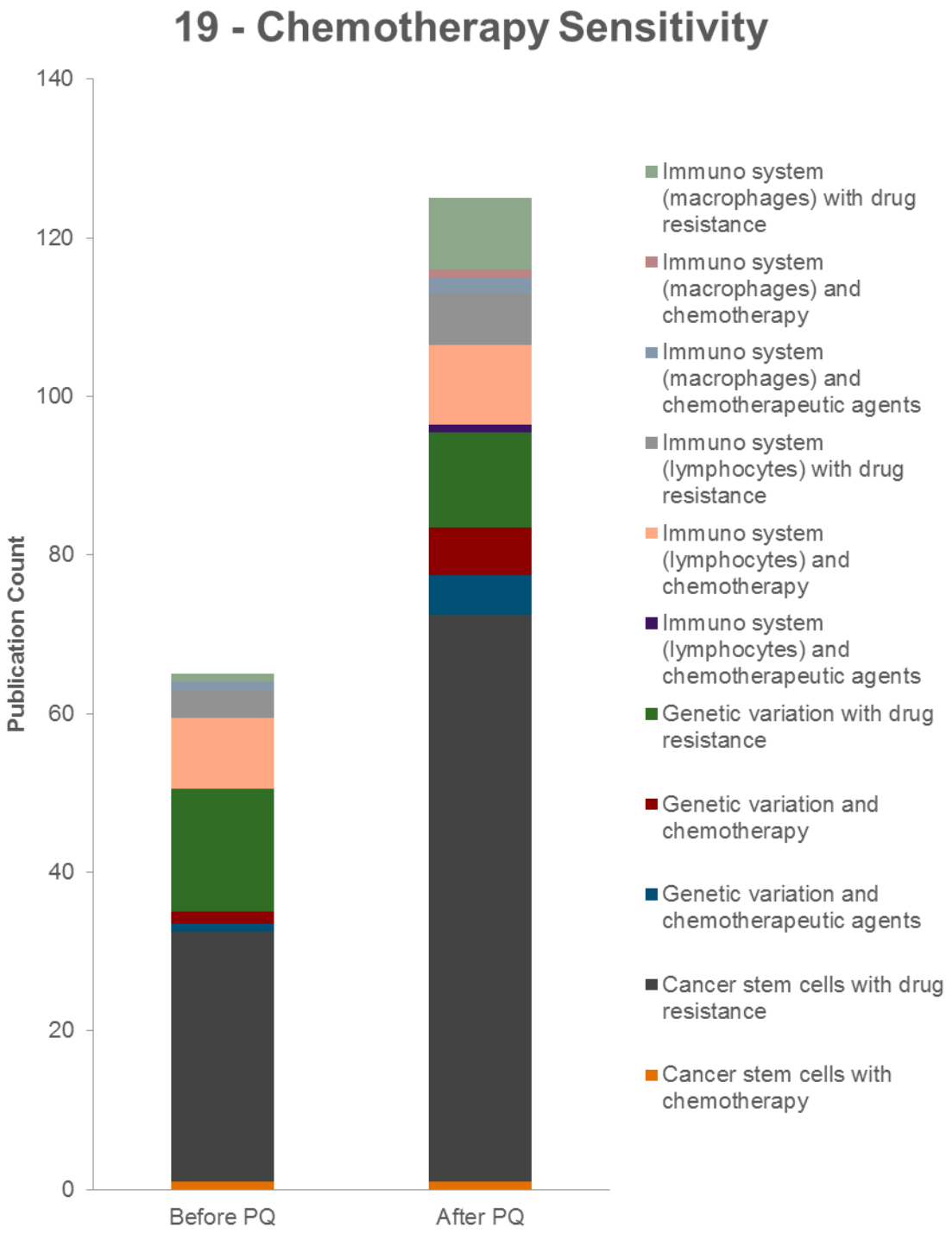


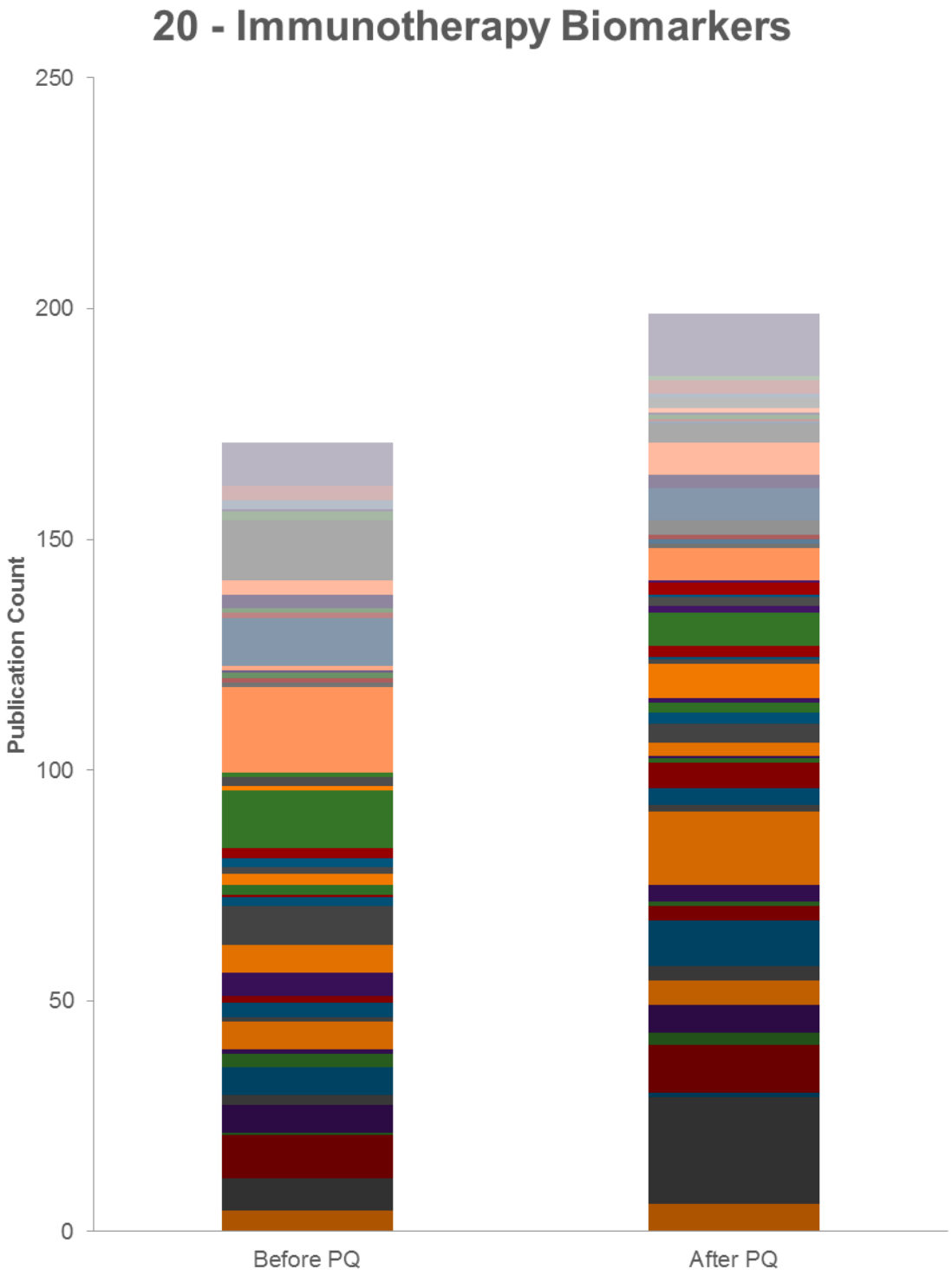
17 - New Drug Testing



18 - Undruggable Targets

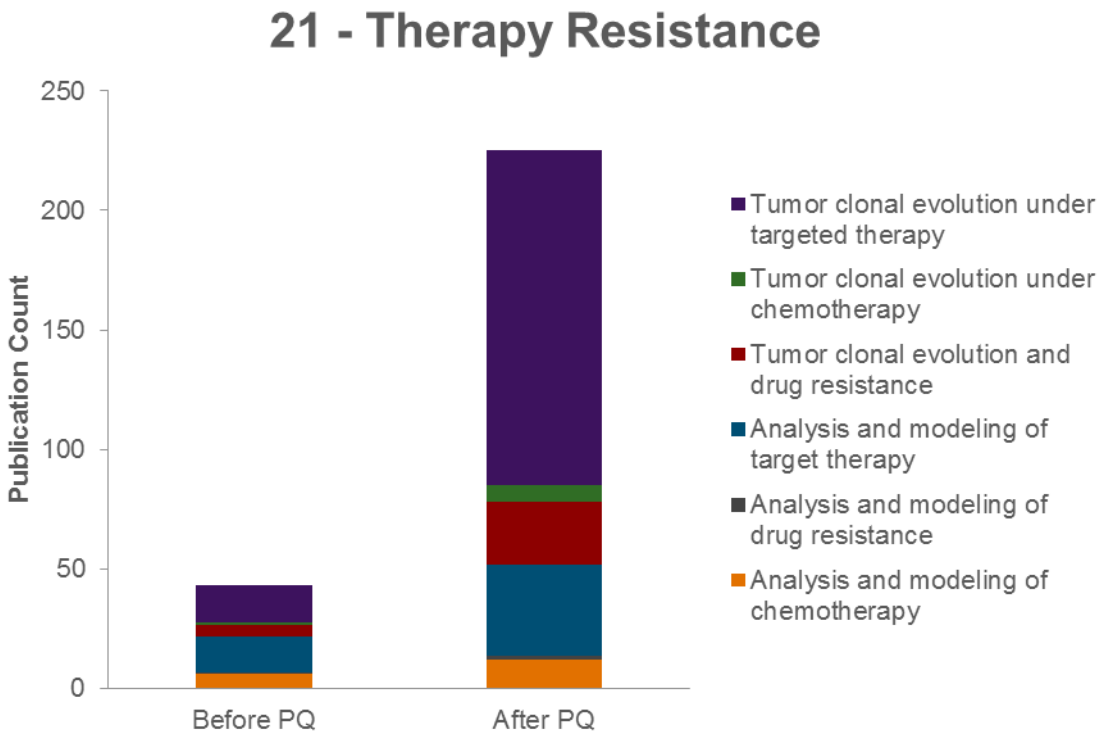



19 - Chemotherapy Sensitivity
Publication Count
140
120
100
80
60
40
20
0
Before PQ
After PQ
Immuno system (macrophages) with drug resistance
Immuno system (macrophages) and chemotherapy
Immuno system (macrophages) and chemotherapeutic agents
Immuno system (lymphocytes) with drug resistance
Immuno system (lymphocytes) and chemotherapy
Immuno system (lymphocytes) and chemotherapeutic agents
Genetic variation with drug resistance
Genetic variation and chemotherapy
Genetic variation and chemotherapeutic agents
Cancer stem cells with drug resistance
Cancer stem cells with chemotherapy

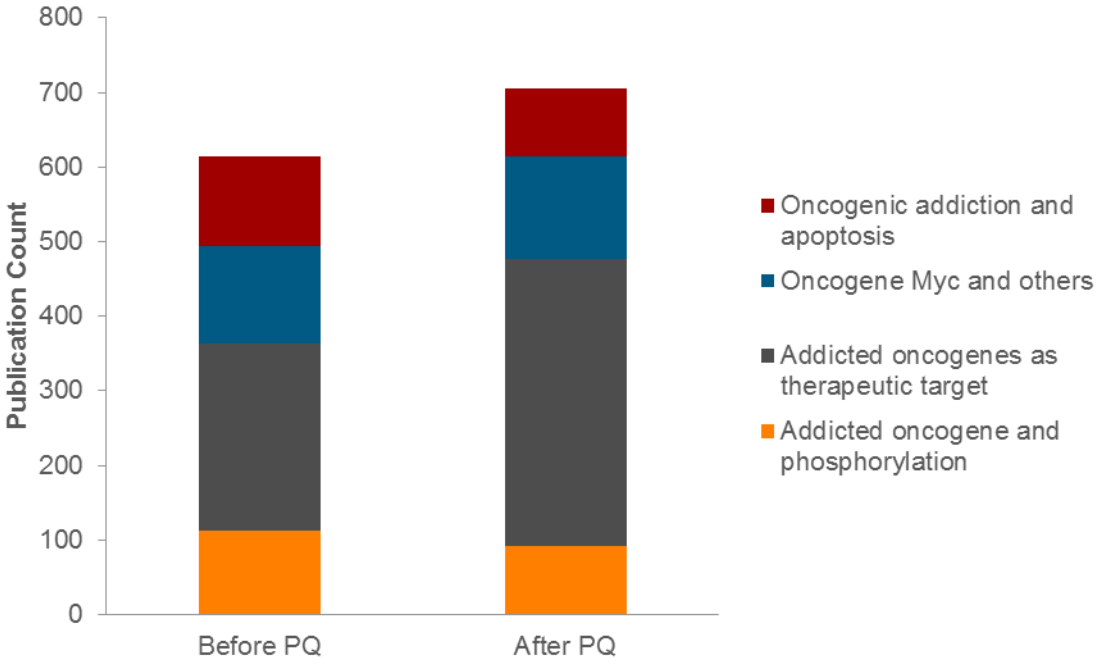

20 - Immunotherapy Biomarkers
250
200
150
100
50
0
Publication Count
Before PQ
After PQ

Legend for PQ 20 – Immunotherapy Biomarkers

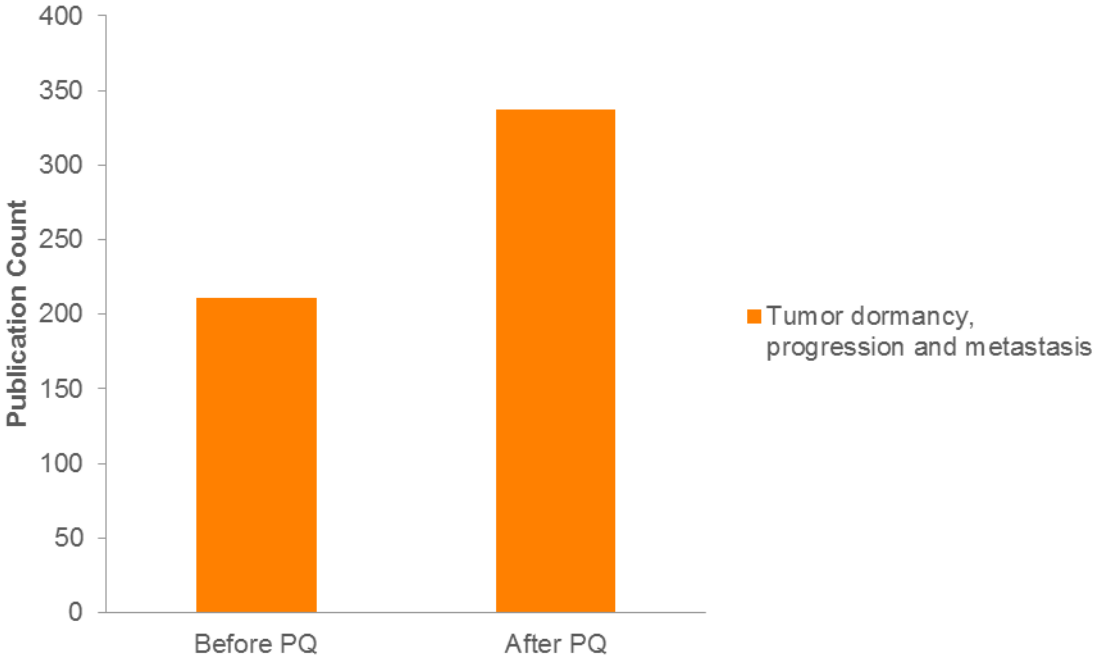
- Tumor biomarkers, cancer and immunoresponse
- Tumor biomarkers for cancer vaccine and immunotherapy
- Tumor biomarkers and tumor infiltrating lymphocyte
- Tumor biomarkers and immunosuppression
- Tumor biomarkers and ILs
- Tumor biomarkers and cancer immunotherapy
- Tumor associated antigens as biomarkers, DC and immunotherapy
- Tumor associated antigens as biomarkers used in Immuno-cell threapy
- Tumor associated antigens as biomarkers and vaccine response
- Tumor associated antigens as biomarkers and proinflammatory cytokines
- Tumor associated antigens as biomarkers and immuno check blockade mAb
- Tumor associated antigens as biomarkers and cancer immunotherapy
- Immunoassay in Immuno-cell therapy
- Immunoassay for DC and immunotherapy
- Immunoassay for vaccine response
- Immunoassay for HSC transplantation
- Immunoassay for cancer vaccine and immunotherapy
- HLA complex and antigen as biomarkers, DC and immunotherapy
- HLA complex and antigen as biomarkers used in Immuno-cell therapy
- HLA complex and antigen as biomarkers and vaccine response
- HLA complex and antigen as biomarkers and NK cells
- HLA complex and antigen as biomarkers and HSC transplantation
- HLA complex and antigen as biomarkers and cancer immunotherapy
- Cell surface receptors as biomarkers, cancer and immunoresponse
- Cell surface receptors as biomarkers for cancer vaccine and immunotherapy
- Cell surface receptors as biomarkers and tumor infiltrating lymphocyte
- Cell surface receptors as biomarkers and immunosuppression
- Cell surface receptors as biomarkers and ILs
- Cell surface receptors as biomarkers and effector T cells
- Cell surface receptors as biomarkers and cancer immunotherapy
- Tumor biomarkers used in Immuno-cell therapy
- Tumor biomarkers and vaccine response
- Tumor biomarkers and NK cells
- Tumor biomarkers and immuno check blockade mAb
- Tumor biomarkers and Checkpoint blocade mAb
- Tumor biomarkers and antigen and cytotoxic lymphocyte
- Tumor associated antigens as biomarkers, cancer and immunoresponse
- Tumor associated antigens as biomarkers for cancer vaccine and immunotherapy
- Tumor associated antigens as biomarkers and tumor infiltrating lymphocyte
- Tumor associated antigens as biomarkers and immunosuppression
- Tumor associated antigens as biomarkers and effector T cells
- Tumor associated antigens as biomarkers and antigen and cytotoxic lymphocyte
- Immunoassay for effector T cells
- Immunoassay for cancer and immunoresponse
- Immunoassay for immunosuppression
- Immunoassay for cytotoxic lymphocyte
- Immunoassay for cancer immunotherapy
- HLA complex and antigen as biomarkers, cancer and immunoresponse
- HLA complex and antigen as biomarkers for cancer vaccine and immunotherapy
- HLA complex and antigen as biomarkers and tumor infiltrating lymphocyte
- HLA complex and antigen as biomarkers and immunosuppression
- HLA complex and antigen as biomarkers and effector T cells
- Cell surface receptors as biomarkers, DC and immunotherapy
- Cell surface receptors as biomarkers used in Immuno-cell therapy
- Cell surface receptors as biomarkers and vaccine response
- Cell surface receptors as biomarkers and NK cells
- Cell surface receptors as biomarkers and immuno check blockade mAb
- Cell surface receptors as biomarkers and HSC transplantation
- Cell surface receptors as biomarkers and Checkpoint blockade mAb
- Cell surface receptors as biomarkers and antigen and cytotoxic lymphocyte



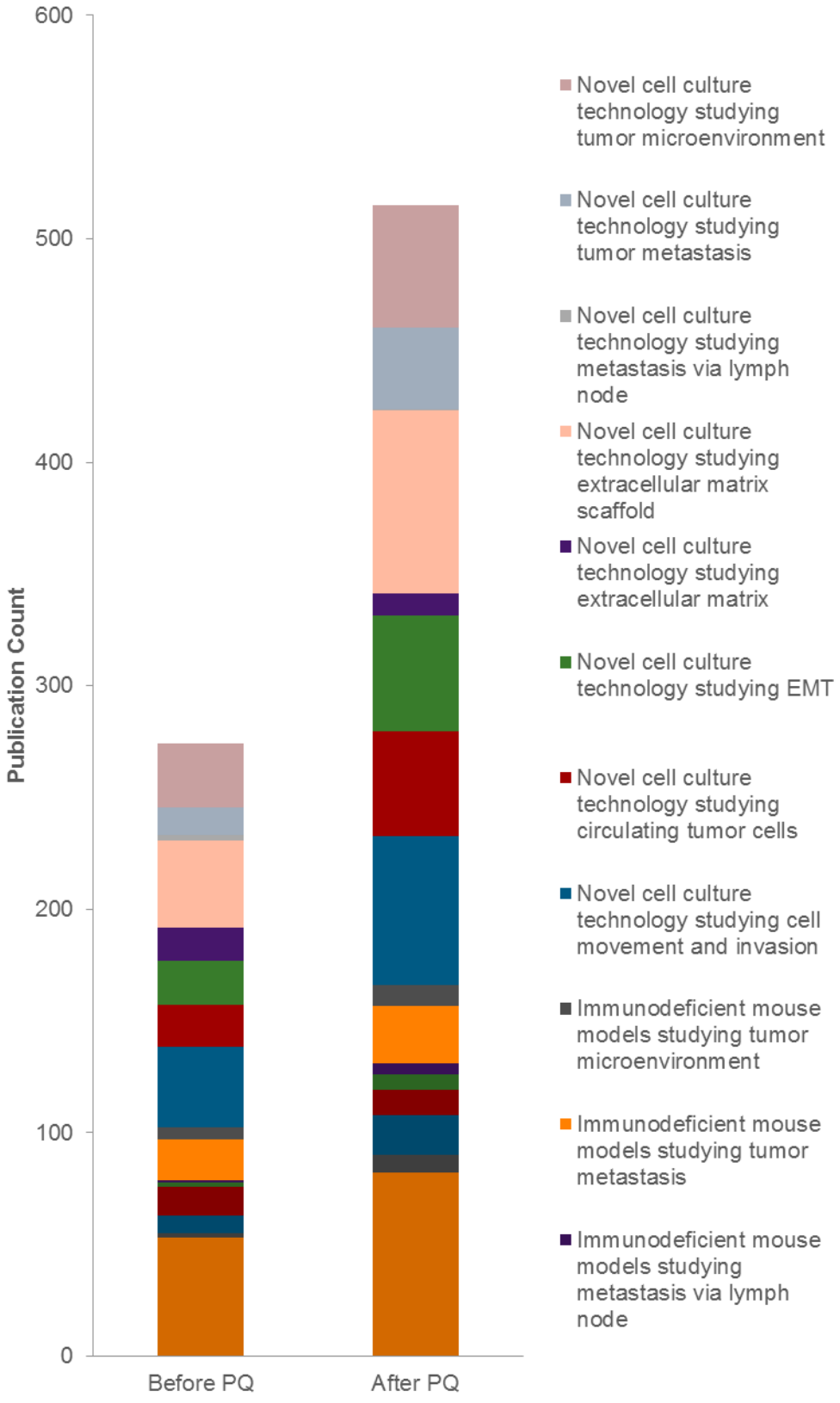
22 - Oncogene Addiction

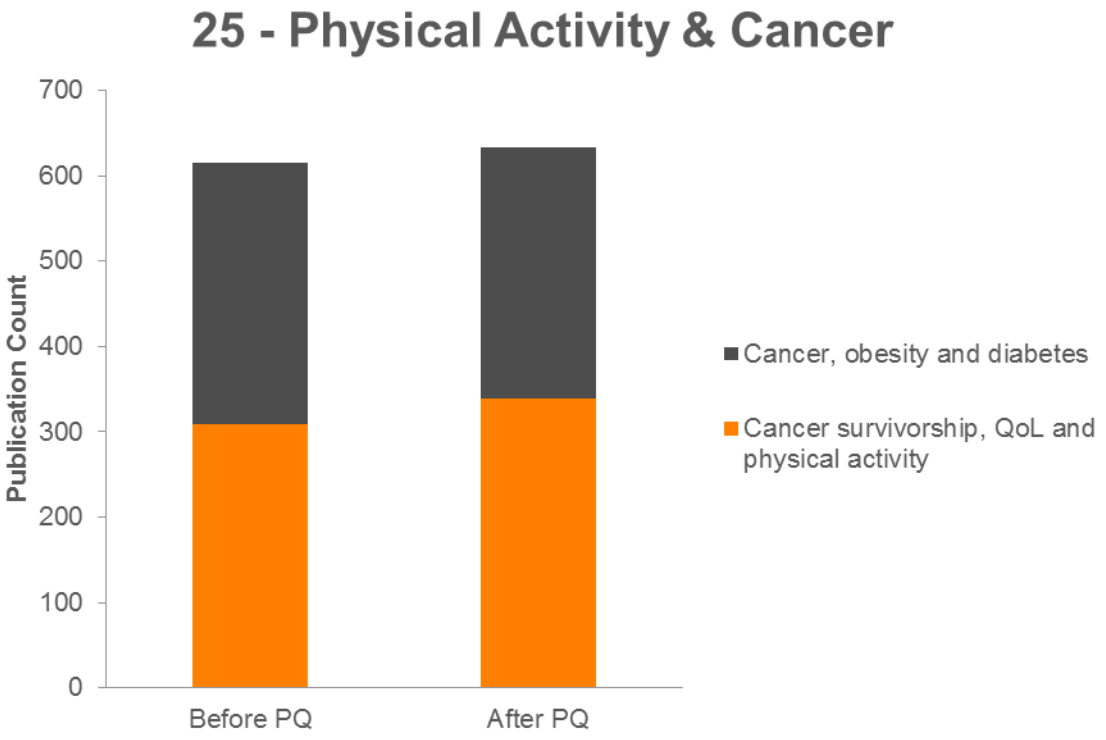


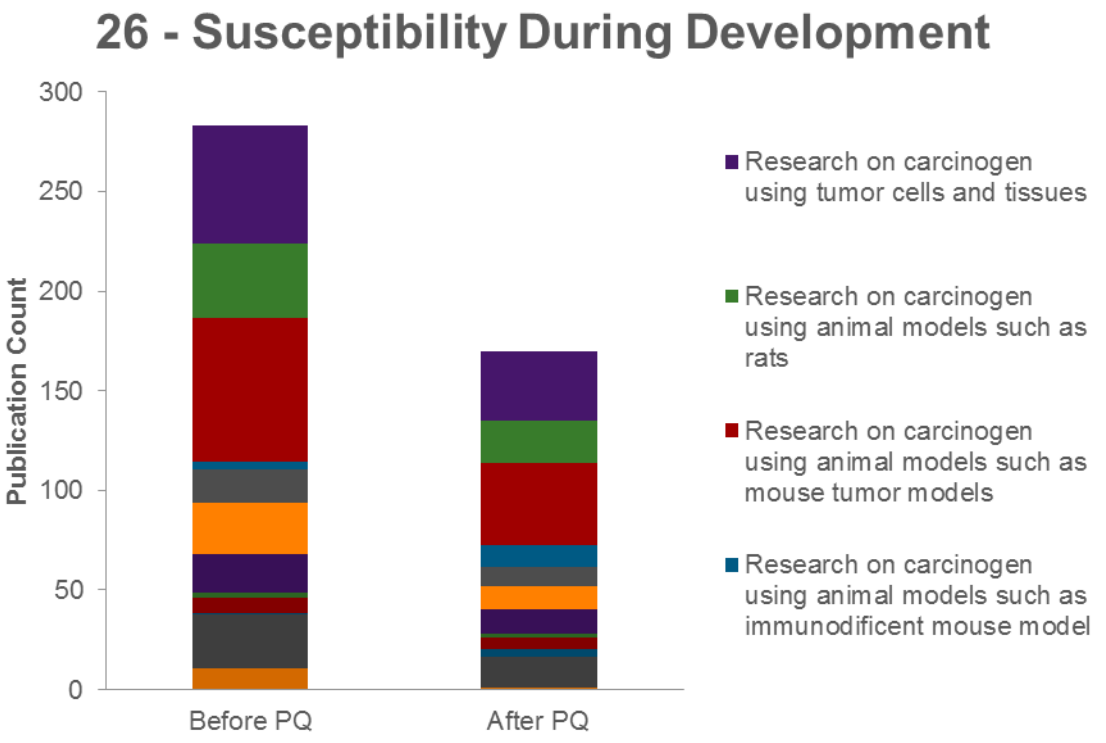
23 - Tumor Indolence

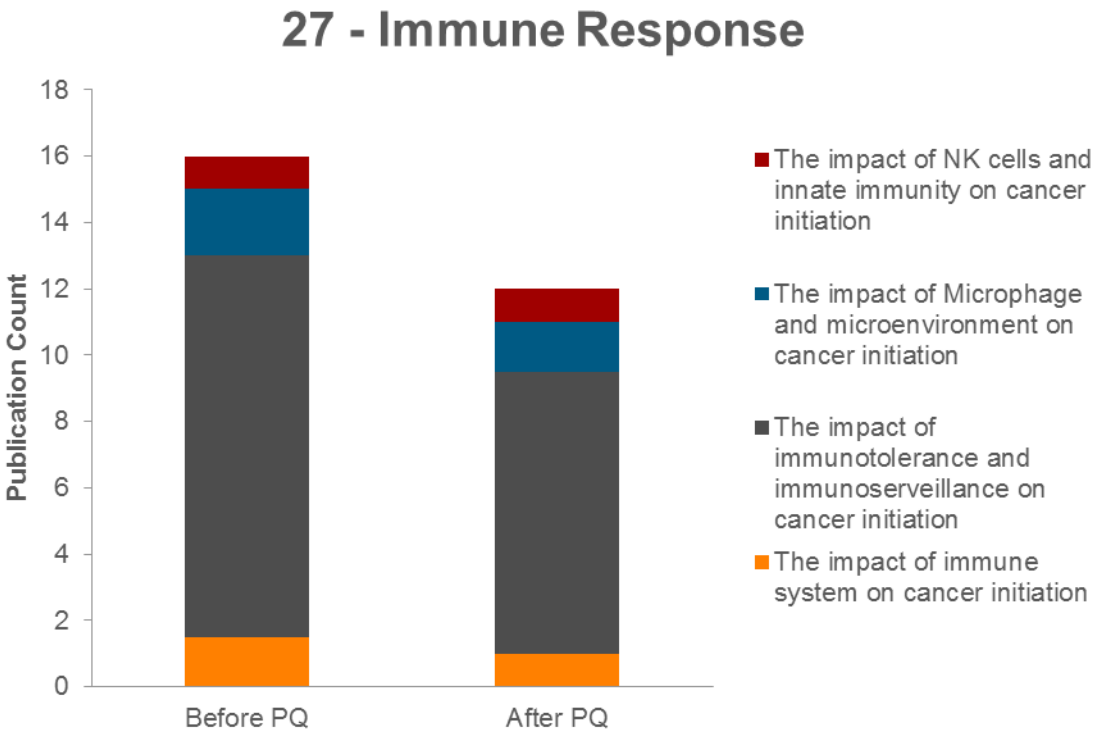


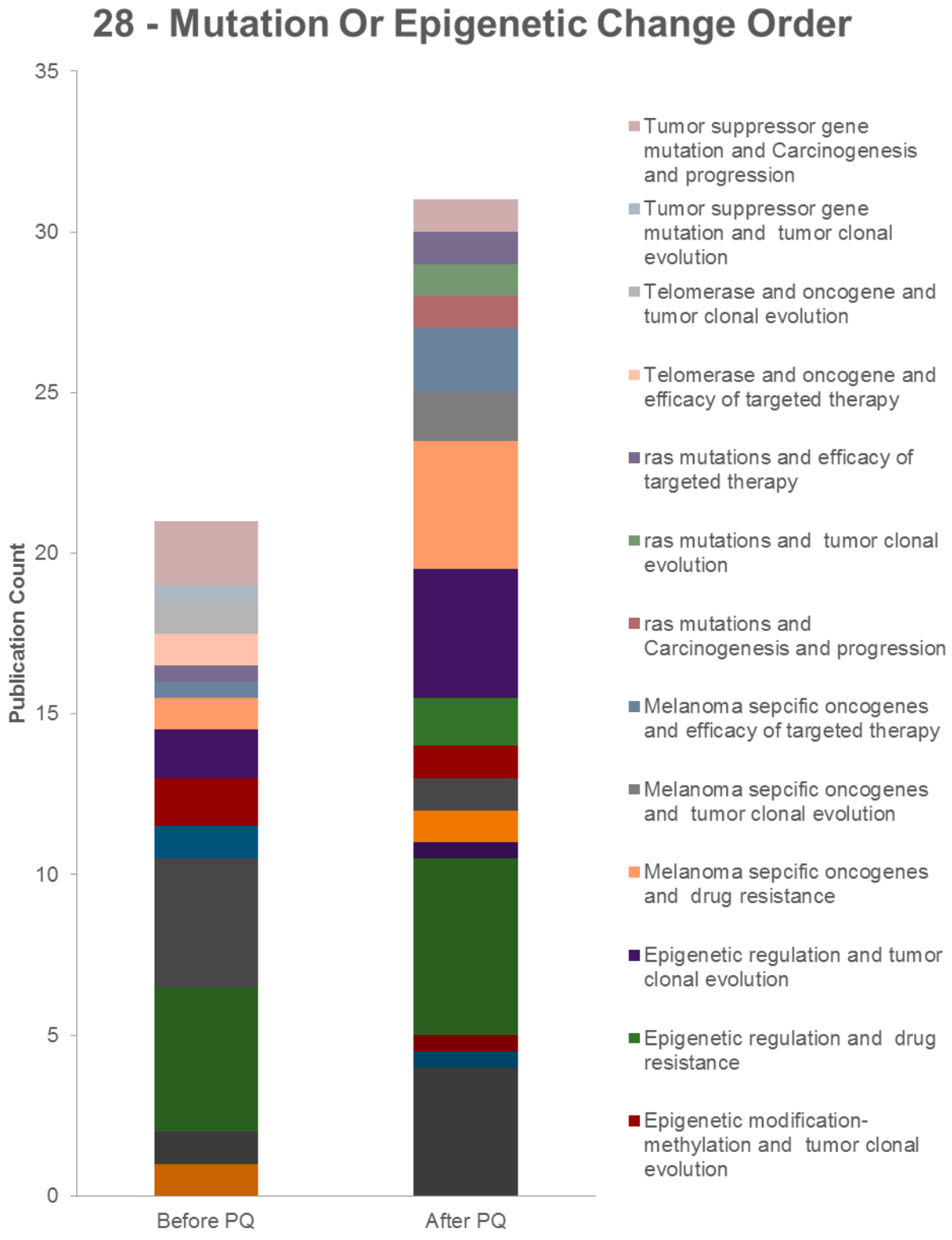
24 - Metastasis Study Techniques



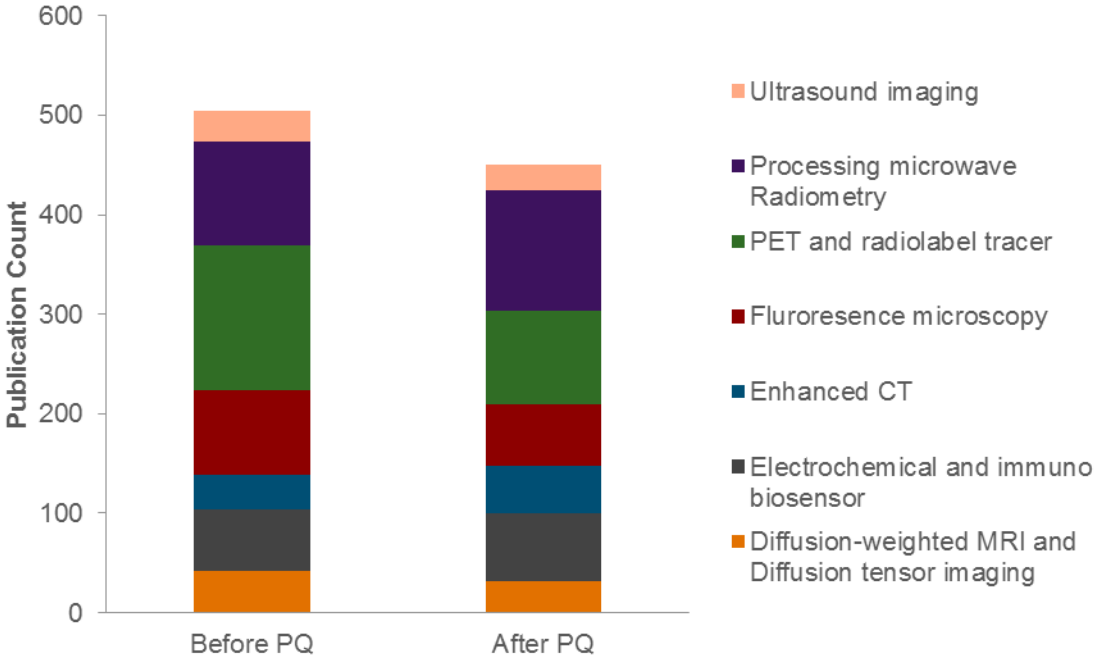

25 - Physical Activity & Cancer
Publication Count
700
600
500
400
300
200
100
0
Before PQ
After PQ
Cancer, obesity and diabetes
Cancer survivorship, QoL and physical activity


26 - Susceptibility During Development
Publication Count
300
250
200
150
100
50
0
Before PQ
After PQ
Research on carcinogen using tumor cells and tissues
Research on carcinogen using animal models such as rats
Research on carcinogen using animal models such as mouse tumor models
Research on carcinogen using animal models such as immunodificent mouse model

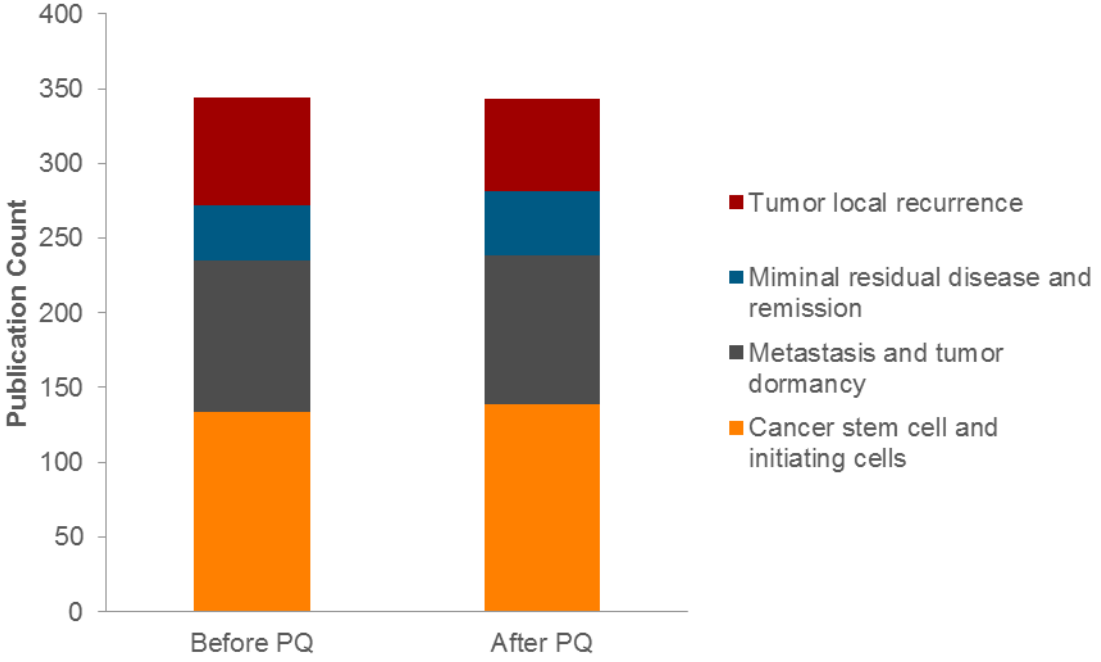

27 - Immune Response
Publication Count
18
16
14
12
10
8
6
4
2
0
Before PQ
After PQ
The impact of NK cells and innate immunity on cancer initiation
The impact of Microphage and microenvironment on cancer initiation
The impact of immunotolerance and immunoserveillance on cancer initiation
The impact of immune system on cancer initiation

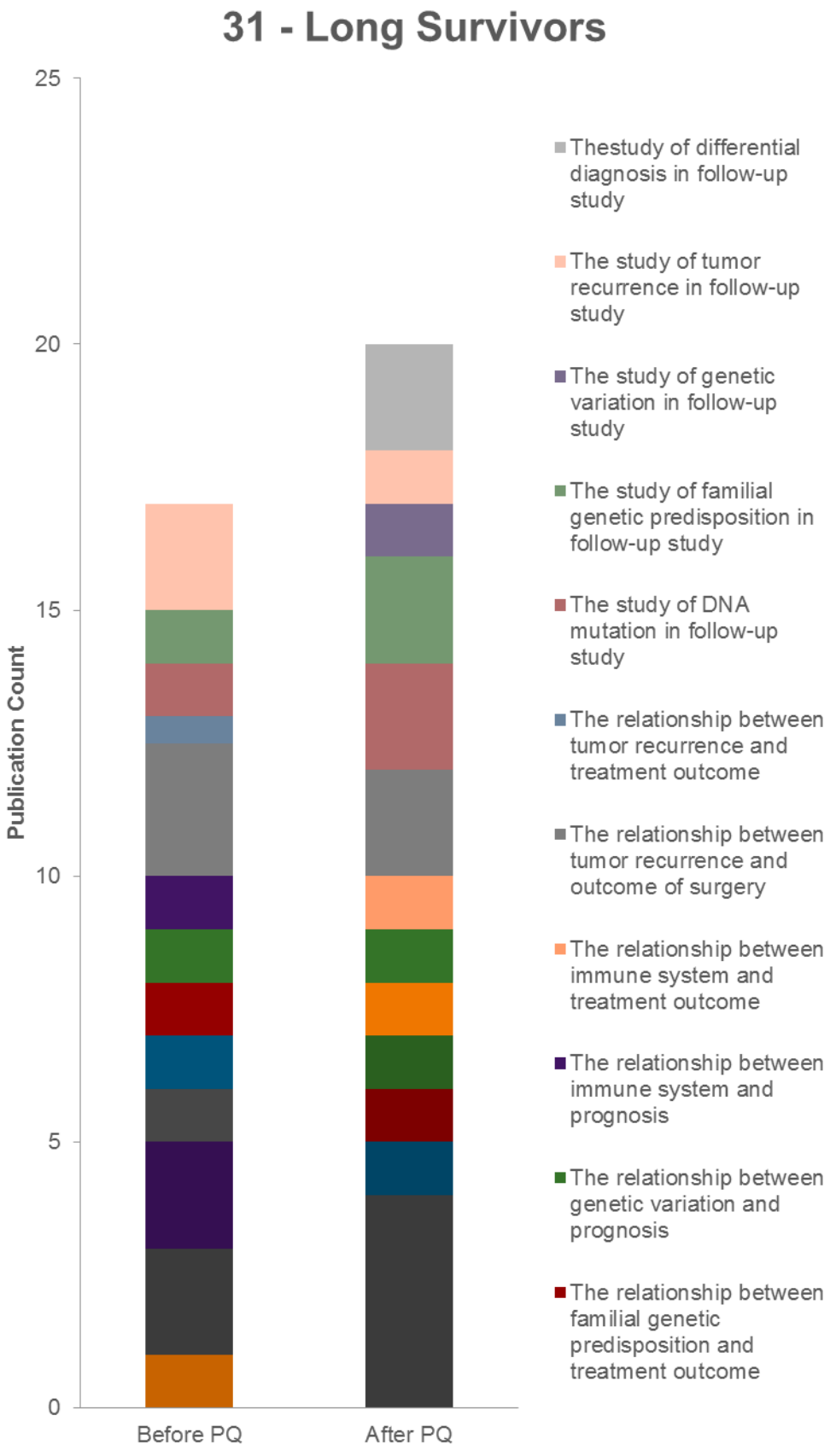

28 - Mutation Or Epigenetic Change Order
Publication Count
0
5
10
15
20
25
30
35
Before PQ
After PQ
Tumor suppressor gene mutation and Carcinogenesis and progression
Tumor suppressor gene mutation and tumor clonal evolution
Telomerase and oncogene and tumor clonal evolution
Telomerase and oncogene and efficacy of targeted therapy
ras mutations and efficacy of targeted therapy
ras mutations and tumor clonal evolution
ras mutations and Carcinogenesis and progression
Melanoma sepcific oncogenes and efficacy of targeted therapy
Melanoma sepcific oncogenes and tumor clonal evolution
Melanoma sepcific oncogenes and drug resistance
Epigenetic regulation and tumor clonal evolution
Epigenetic regulation and drug resistance
Epigenetic modification-methylation and tumor clonal evolution

29 - Physical Properties

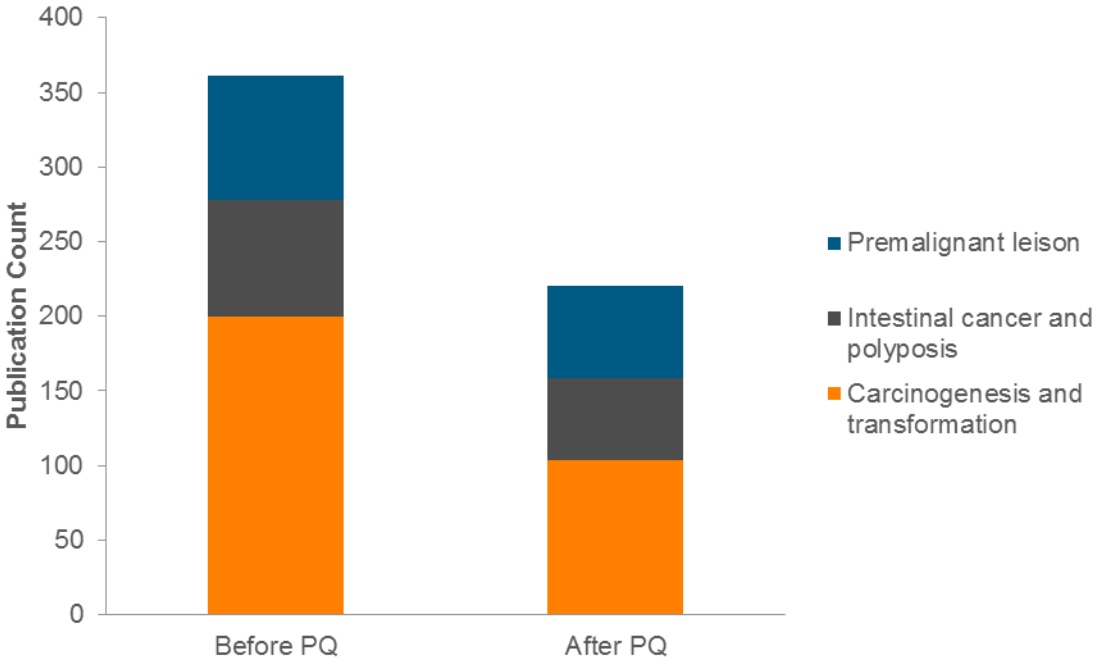


30 - Dormancy And Recurrence




31 - Long Survivors
Publication Count
25
20
15
10
5
0
Before PQ
After PQ
Thestudy of differential diagnosis in follow-up study
The study of tumor recurrence in follow-up study
The study of genetic variation in follow-up study
The study of familial genetic predisposition in follow-up study
The study of DNA mutation in follow-up study
The relationship between tumor recurrence and treatment outcome
The relationship between tumor recurrence and outcome of surgery
The relationship between immune system and treatment outcome
The relationship between immune system and prognosis
The relationship between genetic variation and prognosis
The relationship between familial genetic predisposition and treatment outcome

32 - Cancer Field Effect



33 - Cachexia